

```

QY 1919 GGCTCTTTAGTCCCTAATAGCTTATGGAATGCAATTTATAGATATTAGCTCTCTCCATTAT 1978
Db 675 AlaSerLeuValProAsnSerLeuTyrGlySerIleLeuAspIleArgSerAlaHisSer 694
QY 1979 CTTATGGAGACTGCAACAGAGGTTCAGGAGACCGTGTCTTTTGGTGTGCTGGATTAA 2038
Db 695 AlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeuTyrValSerGlyVal 714
QY 2039 TCTAACTTCTTCCATAGGATAGTACAAAACACGACGCGGTTTCGCCATTTGAGTGGC 2098
Db 715 SerAsnPhePheTyrHisAspArgAlaLeuGlyGlnGlyTyrArgTyrIleSerGly 734
QY 2099 GGTATGTCATAGGAGGAACCTACATCTTGTTCAGATAAGATTCTTACGTCTGCAATT 2158
Db 735 GlyTyrSerLeuGlyAlaAsnSerTyrPheGlySer--SerMetPheGlyLeuAlaPhe 753
QY 2159 TGTACGCTCTTGGAGAGATAGAGACTTCTTGTAGCTAAGAAATCAAGGTACAGTCTAC 2218
Db 754 ThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsnHisAlaCys 773
QY 2219 GGAGGAACCTCTCTATTACAGCACAAACGAAACCTATATCTCTCTCTGCAAACTACGG 2278
Db 774 IleGlySerValTyrLeuSerThrGlnGlnAlaLeu----- 785
QY 2279 CTTGTTCTGTTGCTTATGTTCTTACAGAGATTCTGTTCTCTTTTCAGGA----- 2329
Db 786 ---CysGly--SerTyr-----LeuPheGlyAspAlaPheIle 796
QY 2330 AACCTTAGCTACACCCATACGATAACGATCTGAAACCAAGTATACAAATATCTCTACT 2389
Db 797 ArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGlu 816
QY 2390 GTTAAAGGAAGCTGGGGGAATGATAGTTTCGCTTTAGAAATTCGGTGAAGAGCTCCGATT 2449
Db 817 SerAspValArgTyrAspAsnAsnCysLeuAlaGlyGluIleGlyAlaGlyLeuProIle 836
QY 2450 TGCTTAGATGAAGTGCTCTATT--GAGCAGTACATGCCCTTCATGAAATTCAGTCTT 2506
Db 837 ValIleThrProSerLysLeuTyrLeuAsnGlnLeuArgProPheValGlnAlaGluPhe 856
QY 2507 GTCTATGCACATCAGGAAGGTTTAAAGACAGGACAGAAAGCTCGTGAATTTGGAAGT 2566
Db 857 SerTyrAlaAspHisGluSerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSer 876
QY 2567 ACCCGTCTGTGAATCTTCCTTACCTATCGGATCCGATTTGATAGGAATCAGACTGC 2626
Db 877 GlyHisLeuLeuAsnLeuSerValProValGlyValLysPheAspArgCysSerSerThr 896
QY 2627 CAAGATGCAACGTACAATCTAACTCTGTGTTATCTGTGATCTTGTTCGTAGTAACCCC 2686
Db 897 HisProAsnLysTyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSer 916
QY 2687 GACTGTACGACACACACTCGGAATTAGCGGTGATTTCTGAAAACCTTCGGTACGAATTG 2746
Db 917 GlyThrGluThrThrLeuLeuSerHisGlnGluThrTyrThrThrAspAlaPheHisLeu 936
QY 2747 GCAAGACAAGCTTTAGTCTCTGTCGAGGGAACCAATTTTGGCTTTAACTCAATTTTGA 2806
Db 937 AlaArgHisGlyValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGlu 956
QY 2807 GCCTTTAGCAATTTCTTTGAATTCGCTGGGTGCATCTCGCAATTACAATGTAGACTTA 2866
Db 957 ValTyrGlyHisGlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAla 976
QY 2867 GGAGCAAAATACCAATTC 2884
Db 977 GlySerLysVal***Phe 982

```

Search completed: December 16, 2003, 10:34:02
Job time : 99 secs

Alignment Scores:

Pred. No.: 0 Length: 928
 Score: 4774.00 Matches: 928
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 88.26% Indels: 0
 DB: 12 Gaps: 0

US-09-428-122-1 (1-3000) x US-09-428-122-2 (1-928)

QY 101 ATGAAGTCTCTTCCCTCCCAAGTTGTTATTTCTACATTTGCTATTTCCCTTTCTCTATG 160
 Db 1 MetLysSerSerPheProLysPheValPheSerThrPheAlaIlePheProLeuSerMet 20
 QY 161 ATTGCTACGACGACAGTTTTCGATTCAGTCGAGTTCGATGGGAATAAAATGGTAAT 220
 Db 21 IleAlaThrGluThrValLeuAspSerSerAlaSerPheAspGlyAsnLysAsnGlyAsn 40
 QY 221 TTTTCAGTTCTGAGAGTCAGGAAGATGCTGGAACCTACCTATTTTAAAGGAAATGTC 280
 Db 41 PheSerValArgGluSerGlnGluAspAlaGlyThrThrTyrLeuPheLysGlyAsnVal 60
 QY 281 ACTCTAGAAAATATCTCTGGAACGACGACGAAATCAGAAAAGCTGTTTAAACAACCT 340
 Db 61 ThrLeuGluAsnIleProGlyThrGlyThrAlaIleThrLysSerCysPheAsnAsnThr 80
 QY 341 AAGGGCGATTTCAGCTTTCACAGGTAAACGGGAACCTCTATTGTTCCAAACGGTGGATGCA 400
 Db 81 LysGlyAspLeuThrPheThrGlyAsnGlyAsnSerLeuLeuPheGlnThrValAspAla 100
 QY 401 GGGAGTGTAGCGGGCTCTGTTTAAACAGCAGCGGTGTAGATAATCTACACAGTTTATA 460
 Db 101 GlyThrValAlaGlyAlaAlaValAsnSerSerValValAspLysSerThrThrPheIle 120
 QY 461 GGGTTTCTTCGCTATCTTTTATGCTCTCTCGAAGTTCGATAACTACCGGCAAGGA 520
 Db 121 GlyPheSerSerLeuSerPheIleAlaSerProGlySerSerIleThrThrGlyLysGly 140
 QY 521 GCGTTAGTCTCTACGGGTAGCTTGTGTTTGACAAAATGTCTCAGTTGCTCTCTCAGC 580
 Db 141 AlaValSerCysSerThrGlySerLeuSerLeuThrLysAsnValSerLeuLeuPheSer 160
 QY 581 AAAAATCTTTCACGATATGCGGTGCTATCACCGCAAAACCTTTTCAATTAACAGGG 640
 Db 161 LysAsnPheSerThrAspAsnGlyGlyAlaIleThrAlaLysThrLeuSerLeuThrGly 180
 QY 641 ACTACAATGCTAGCTGTTTCTGAAAATACCTCTCTCAAGAAAGCGGACCAATCAG 700
 Db 181 ThrThrMetSerAlaLeuPheSerGluAsnThrSerSerLysLysGlyGlyAlaIleGln 200
 QY 701 ACTTCGATGCCCTTACCATTTACTGGAACCAAGGGAAGTCTCTTTTCTGACAAATCT 760
 Db 201 ThrSerAspAlaLeuThrIleThrGlyAsnGlnGlyGluValSerPheSerAspAsnThr 220
 QY 761 TCTTCGATCTCGAGCTGCAATTTTTCAGAAAGCTCGGTGACTTATTTCTTAATATGCT 820
 Db 221 SerSerAspSerGlyAlaAlaIlePheThrGluAlaSerValThrIleSerAsnAla 240
 QY 821 AAGTTTCTTTTATTCACATAGCTCAGGAGCGGCTCTCAACACGCGGGATATG 880
 Db 241 LysValSerPheIleAspAsnLysValThrGlyAlaSerSerSerThrThrGlyAspMet 260
 QY 881 TCAGGAGGTCTATCTGTGCTTATAAACTAGTACAGATCTAAAGCTCACCCCTCACTGGA 940
 Db 261 SerGlyGlyAlaIleCysAlaTyrLysThrSerThrAspThrLysValThrLeuThrGly 280
 QY 941 AATCAGATGTTACTCTTCAGCAACAATATCATGCAACAGCGGGAGGAGCTATCTATGTG 1000
 Db 281 AsnGlnMetLeuLeuPheSerAsnAsnThrSerThrAlaGlyGlyAlaIleTyrVal 300
 QY 1001 AAAAAGCTCCAACTGCTCCGAGGACTTACCTATTTCAGTAGAAAATAGTCTCAATGGA 1060
 Db 301 LysLysLeuGluLeuAlaSerGlyGlyLeuThrLeuPheSerArgAsnSerValAsnGly 320

QY 1061 GGTACAGCTCTCTAAAGTGGAGCCATAGCTATCGAAGATAGTGGGAATTCAGTTTATCC 1120
 Db 321 GlyThrAlaProLysGlyGlyAlaIleAlaIleGluAspSerGlyGluLeuSerLeuSer 340
 QY 1121 GCGATAGTGTGACATTTCTTTTAGGGAATACAGTCACTTCTTACTACTCTCTGGGACG 1180
 Db 341 AlaAspSerGlyAspIleValPheLeuGlyAsnThrValThrSerThrThrProGlyThr 360
 QY 1181 AATAGAGTACTATCGACTTAGAAGAGTGCAGATGACAGCTTTCGCTCTGCTGCT 1240
 Db 361 AsnArgSerSerIleAspLeuGlyThrSerAlaLysMetThrAlaLeuArgSerAlaAla 380
 QY 1241 GGTAGAGCCATCTACTTCTATGATCCCATAACTACAGATCATCCACAACTAGTTACAGAT 1300
 Db 381 GlyArgAlaIleTyrPheTyrAspProIleThrThrGlySerSerThrThrValThrAsp 400
 QY 1301 GTCCTAAAAGTTAATGAGACTCCGACAGATTCGCACTACAAATATACAGGGAACATATC 1360
 Db 401 ValLeuLysValAsnGluThrProAlaAspSerAlaLeuGlnTyrThrGlyAsnIleIle 420
 QY 1361 TTCACAGGAGAAAAGTTATCAGAGACAGAGCGCGAGATTCTTAAAAATCTTACTTCGAAG 1420
 Db 421 PheThrGlyGluLysLeuSerGluThrGluAlaAlaAspSerLysAsnLeuThrSerLys 440
 QY 1421 CTACTACAGCTGTAACTCTTTTACAGAGTACTCTATCTTTTAAAAACATGGAGTACTCTG 1480
 Db 441 LeuLeuGlnProValThrLeuSerGlyGlyThrLeuSerLeuLysHisGlyValThrLeu 460
 QY 1481 CAGACTCAGGCATTCACCAACAGCAGATTCCTGCTCTCGAAATCGAGTAGGAGTACT 1540
 Db 461 GlnThrGlnAlaPheThrGlnGlnAlaAspSerAlaGluGluMetAspValGlyThrThr 480
 QY 1541 CTAGAACCCTGCTACTAGCACCAATAACAATTTGGTCAATTAACATCAGTTCTTATAGAC 1600
 Db 481 LeuGluProAlaAspThrSerThrIleAsnAsnLeuValIleAsnIleSerSerIleAsp 500
 QY 1601 GGTGCAAGAGCGCAAAATAGAACCAAGCTAGCTCAAAAATCTCAAACTTACTCGA 1660
 Db 501 GlyAlaLysAlaLysIleGluThrLysAlaThrSerLysAsnLeuThrLeuSerGly 520
 QY 1661 ACCATCAGCTTATTTGACCCGACGGGACGCTTTTATGAAATCATAGTTTAAAGAAATCCT 1720
 Db 521 ThrIleThrLeuLeuAspProThrGlyThrPheTyrGluAsnHisSerLeuArgAsnPro 540
 QY 1721 CAGTCTACGACATCTTAGAGCTCAAGCTTCTGAACTGTGAACTGTAAACAGCACCCAGT 1780
 Db 541 GlnSerTyrAspIleLeuGluLeuLysAlaSerGlyThrValThrSerThrAlaValThr 560
 QY 1781 CCAGATCCTATATGCGTGAAGAAATTCATATACGGCTATCAGGAACTTGGGGCCCAAT 1840
 Db 561 ProAspProIleMetGlyGlyLysPheHisTyrGlyTyrGlnGlyThrTrpGlyProIle 580
 QY 1841 GTTTGGGGACAGGGCTCTACAGCTGCAACCTTCACTGCACTGAACTAAAAGTCTATATT 1900
 Db 581 ValTrpGlyThrGlyAlaSerThrThrAlaThrPheAsnTrpThrLysThrGlyTyrIle 600
 QY 1901 CCTAATCCCGAGGCTATCGCTCTTTAGTCCCTATATAGCTTATGGAATGCATTTATAGAT 1960
 Db 601 ProAsnProGluArgIleGlySerLeuValProAsnSerLeuTrpAsnAlaPheIleAsp 620
 QY 1961 ATTAGCTCTCTCATTTATTTTGGAGAGTGCAGAAAGGTTTCAGGAGACCGTGTCT 2020
 Db 621 IleSerSerLeuHisTyrLeuMetGluThrAlaAsnGluGlyLeuGlnGlyAspArgAla 640
 QY 2021 TTTTGTGTGCTGATTTACTTCTTCCATAGGATAGTACAAAACACAGCGCGG 2080
 Db 641 PheTrpCysAlaGlyLeuSerAsnPhePheHisLysAspSerThrLysThrArgGly 660
 QY 2081 TTTCCGCTTGTAGTGGCGTTATGTATGATAGGAGAAACCTCATACTTTGTTTACAGTAAG 2140
 Db 661 PheArgHisLeuSerGlyGlyTyrValIleGlyGlyAsnLeuHisThrCysSerAspLys 680

QY 2141 ATTCTTAGTCTGCTGCTTTTCTAGCTCTTTTGGAGAGATAGAGACTACTTTTGTAGCTAAG 2200
Db 681 IIELEUSERAlaAlaPheCysGlnLeuPheGlyArgAspArgAspPyrPheValAlaLys 700
QY 2201 AATCAGGTACAGTCTACGAGGAACTCTATTACAGCAGACAAACCACTTATCTCT 2260
Db 701 AsnGlnGlyThrValTyrGlyGlyThrLeuTyrTyrGlnHisAsnGluThrTyrIleSer 720
QY 2261 CTTCTCTGCAACCTACCGCTTGTCTGCTCTTATGTTCTCTACAGAGATCTCTGTTCTC 2320
Db 721 LeuProCysLysLeuArgProCysSerLeuSerTyrValProThrGluLeuProValLeu 740
QY 2321 TTTTCAGGAAACCTTAGCTACACCATACCGATACGATCTGAAACCCAGATATACACA 2380
Db 741 PheSerGlyAsnLeuSerTyrThrHisThrAspAsnAspLeuLysThrIleTyrThrThr 760
QY 2381 TATCCTACTCTTAAGGAAGCTGGGGAATGATAGTTTCGCTTAGAATTCGTTGGAAGA 2440
Db 761 TyrProThrValLysGlySerTyrGlyAsnAspSerPheAlaLeuGluPheGlyGlyArg 780
QY 2441 GCTCCGATTTGCTTAGATGAAAGTCTCTATTGAGCAGTACATGCGCTTTCATGAAATTC 2500
Db 781 AlaProIleCysLeuAspGluSerAlaLeuPheGluGlnTyrMetProPheMetLysLeu 800
QY 2501 CAGTTTCTATGCATCAGGAAGTTTAAAGACAGGGAACAGAGCTCTGTAATTT 2560
Db 801 GlnPheValTyrAlaHisGlnGluGlyPheLysGluGlnGlyThrGluAlaArgGluPhe 820
QY 2561 GGAAGTAGCGCTTGTGAACTTGTCTTACCTATCGGATCCGATTGTAAGGAATCA 2620
Db 821 GlySerSerArgLeuValAsnLeuAlaLeuProIleGlyLeuArgPheAspLysGluSer 840
QY 2621 GACTGCCAAGATGCAAGTCAATCTAACTCTGTTGTTATCTGTGATCTGTTCTGATG 2680
Db 841 AspCysGlnAspAlaThrTyrAsnLeuThrLeuGlyTyrThrValAspLeuValArgSer 860
QY 2681 AACCCGACTGTACGACCACTCGGAATTAGCGGTGATCTTGGAAACCTTCGCTAGC 2740
Db 861 AsnProAspCysThrThrThrLeuArgIleSerGlyAspSerTyrPheGlyThr 880
QY 2741 AATTGGCAAGACAAAGCTTTAGTCTCTGCTGCGGGAACCAATTTTGTCTTAACTCAAAT 2800
Db 881 AsnLeuAlaArgGlnAlaLeuValLeuArgAlaGlyAsnHisPheCysPheAsnSerAsn 900
QY 2801 TTTCAGCCCTTACCAATTTCTTTTGAATTCGCTGCGGTCATCTCGCAATTCAGATGA 2860
Db 901 PheGluAlaPheSerGlnPheSerPheGluLeuArgGlySerSerArgAsnTyrAsnVal 920
QY 2861 GACTTAGGAGCAAAATACCAATTC 2884
Db 921 AspLeuGlyAlaLysTyrGlnPhe 928

RESULT 2

US-09-452-380-3
; Sequence 3, Application US/09452380
; Patent No. US20020094340A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, Andrew D.
; APPLICANT: COMEN, Raymond P.
; APPLICANT: WANG, Joe
; APPLICANT: DUNN, Pamela
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE REFERENCE: 032931/0216
; CURRENT APPLICATION NUMBER: US/09/452,380
; CURRENT FILING DATE: 1999-12-01
; PRIORITY APPLICATION NUMBER: US 60/132,272
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: US 60/113,439
; PRIOR FILING DATE: 1998-12-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3

; LENGTH: 936
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-452-380-3
Alignment Scores:
Pred. No.: 6.38e-153 Length: 936
Score: 1965.00 Matches: 411
Percent Similarity: 62.30% Conservative: 179
Best Local Similarity: 43.40% Mismatches: 327
Query Match: 36.33% Indels: 30
DB: 9 Gaps: 16
US-09-428-122-1 (1-3000) x US-09-452-380-3 (1-936)
QY 101 ATGAAGTCTCTTCCCAAGTTTGTATTTCTACATTTGCTATTTC---CCTTTGCT 157
Db 1 MetLysSerSerValSerTrpLeuPhePheSerSerIleProLeuPheSerSerLeuSer 20
QY 158 ATGATTGCTACCGAGACAGTTTTCGATTCAAGT---GCGAGTTTCGATGGGAATATAAAT 214
Db 21 IleValAlaAlaGluValThrLeuAspSerSerAsnAsnSerTyrAspGlySerAsnGly 40
QY 215 GGTAAATTTTCAGTTTCGTGAGAGTCAGGAA---GATGCTGGAACCTACCTATTATTAAG 271
Db 41 ThrThrPheThrValPheSerThrThrAspAlaAlaGlyThrThrThrThrThrThr 60
QY 272 GGAATGTCACTCTAGAAAATATTTCCTGGAAACAGGCACAGCAATCACAAAAGCTGTTTT 331
Db 61 SerAspValSerPheGlnAsnAlaGlyAlaLeuGlyIleProLeuAlaSerGlyCysPhe 80
QY 332 AACACACTAAGGCGATTTCACCTTTCACAGTACCGGAACCTCTATTGTTCCAAACG 391
Db 81 LeuGluAlaGlyGlyAspLeuThrPheGlnGlyAsnGlnHisAlaLeuLysPheAlaPhe 100
QY 392 GTGAGTACGAGGACCTGTAGCAGCGGCTCTCTTAACACAGCGGTGTAGATAAATCTACC 451
Db 101 IleAsnAlaGlySerSerAlaGlyThrValAlaSerThrSerAlaAlaAspLysAsnLeu 120
QY 452 AGTTTATAGGTTTCTCTCTCTATCTTTTATTCGTTCTCTGGAAGTTCGATAACT--- 508
Db 121 LeuPheAsnAspPheSerArgLeuSerIleIleSerCysProSerLeuLeuLeuSerPro 140
QY 509 ACCGCAAGGAGCGTTAGCTGTCTACGCGTAGCTTACAGTGTAGTTGACAAAAGTCACT 568
Db 141 ThrGlyGlnCysAlaLeuLys---SerValGlyAsnLeuSerLeuThrGlyAsnSerGln 159
QY 569 TTGCTCTTCAGCAAAACCTTTTCAACGGAATAATGCGGTGCTATCACCGCAAAACCTCTT 628
Db 160 IleIlePheThrGlnAsnPheSerSerAspAsnGlyGlyValIleAsnThrLysAsnPhe 179
QY 629 TCATTAAACAGGACTACAAATGTACGCTCTCTGTTTCTGAAATACG-----TCCTCAAG 682
Db 180 LeuLeuSerGlyThrSerGlnPheAlaSerPheSerArgAsnGlnAlaPheThrGlyLys 199
QY 683 AAAGCGGAGCGCATTCACACTTCCGATCCCTTACCATTTACTGGAAACCAAGGGGAGTC 742
Db 200 GlnGlyGlyValValTyrAlaThrGlyThrIleThrIleGluAsnSerProGlyIleVal 219
QY 743 TCTTTTCTGACAATCTCTCTTCGGAATCTCGAGCTGCAATTTTACAGAGACCTCGGTG 802
Db 220 SerPheSerGlnAsnLeuAlaLysGlySerGlyGlyAlaLeuTyrSerThrAspAsnCys 239
QY 803 ACTATTCTTAATGCTAAAGTTTCTCTTATTGACATAAGGTCACAGGAGCGGAGCTCC 862
Db 240 SerIleThrAspAsnPheGlnValIlePheAspGlyAsnSerAlaTrpGluAlaGln 259
QY 863 TCAACAACGGGGATATCTCAGAGGCTCTATCTGTGCTTATATAAACTAGTACAGATCT 922
Db 260 Ala-----GlnGlyGlyAlaIleCysCys-----ThrThrThrAspLys 272
QY 923 AAGGTCACCTCCTGGAATTCAGATGTTTACTTTCAGCAACAATATACATCCAGCAACGCG 982

273	ThrVal	ThrLeu	ThrGly	Asn	Leu	Ser	Phe	Thr	Asn	Asn	Thr	Ala	Leu	Thr	Tyr	292
983	GGAGGAG	CATCTATCTATGTG	AAAAAGCTCGAACTGGCTTCCGGAGG	AGCTTACCCCTATT	CAGT	1042										
293	GlyGly	Ala	Ile	Ser	Gly	Leu	Val	Ser	Ile	Ser	Ala	Gly	Gly	Pro	Thr	Leu
1043	AGAAAT	AGTGTCTAAATGG	AGGTACAGCTCTCAA	-----GGTGG	AGCCATAGCTATCG	1096										
313	Ser	Asn	---Ile	Ser	Gly	Ser	Ser	Ala	Gly	Gln	Gly	Gly	Ala	Ile	Asn	Ile
1097	GATAGTGGG	AAATTGAGTTTATCCGCGA	TAGTGGTGACATGCTCTCTTTT	TAGGGAATACA	1156											
332	Ser	Ala	Gly	Glu	Leu	Ala	Leu	Ser	Ala	Thr	Ser	Gly	Asp	Ile	Thr	Phe
1157	GTCACTTCT	ACTACTCTCTGGGACA	ATAGAAATAGTATGACATTAGGAA	CGACGTGCA	1216											
352	Val	Thr	Asn	Gly	Ser	Thr	---Arg	Asn	Ala	Ile	Asn	Ile	Ile	Asp	Thr	Ala
1217	ATCAGACG	TTTGGTCTGCTGCTGGT	PAGAGCATTCTCTATGATCC	ATAACTACA	1276											
371	Val	Thr	Ser	Ile	Arg	Ala	Ala	Thr	Gly	Gln	Ser	Ile	Thr	Phe	Thr	Asn
1277	GGATCATCC	ACAACAGTTACAGATGCTT	AAAAAGTTAATGAGACTCCG	CAGATTCTGCA	1336											
391	Pro	Gly	Thr	Ala	Ala	Ser	Thr	Leu	Asn	Leu	Ala	Asp	Ala	Asn	Ser	Glu
1337	CTCAATAT	CACAGGACATCATCTTCA	CAGGAGAAAGTTATCAGACAG	CAGAGGCGCA	1396											
411	Ile	Glu	Thr	Gly	Ala	Ile	Val	Phe	Ser	Gly	Glu	Leu	Ser	Pro	Thr	Gly
1397	GATTCTAAA	AATCTTACTTCCGAAGCT	ACTACAGCCTCTAACTCTTTC	CAGAGGTACTCTA	1456											
431	Ile	Ala	Ala	Asn	Val	Thr	Ser	Thr	Ile	Arg	Gln	Pro	Ala	Val	Leu	Ala
1457	TCTTTAAA	ACATGGAGTACTCTCGAC	ACTCAGCATTCACTCA	CAGGACAGATTCTCGT	1516											
451	Val	Leu	Arg	Asp	Gly	Val	Thr	Val	Thr	Phe	Leu	Asp	Leu	Thr	Gln	Ser
1517	CTCGAAAT	GACGTAGGAACCTACTCT	AGAACCTGCTCATCTAGC	---ACCATAA	1573											
471	Ile	Leu	Met	Asp	Gly	Gly	Thr	Thr	Leu	Ser	Ala	Gly	Ala	Asn	Leu	Ser
1574	TTGTCTAAT	AATCAATCAGTCTATAG	ACGGTGCAAAGAGCAAAAT	ATAGAAACCA	1633											
491	Leu	Ala	Val	Asn	Leu	Ser	Ser	Leu	Asp	Gly	Thr	Asn	Leu	Ala	Leu	Thr
1634	ACGTCAAA	AATCTGACTTATCTGGA	ACCATCACTTTATTGGAC	CGACGGCAGCTTT	1693											
511	Ala	Asp	Leu	Asn	Ile	Ser	Leu	Ser	Gly	Thr	Ile	Ala	Leu	Ile	Asp	Thr
1694	TATGAAAT	CATAGTTTATAGAAAT	PCCTCAGTCCTACGACATCTT	PAGAGCTCAA	1747											
531	Tyr	Glu	Asn	His	Asn	Leu	Leu	Ser	Ala	Ser	Thr	Tyr	Pro	Leu	Leu	Glu
1748	---GCTTCT	CGAACTGTAAACAGC	ACCGAGTCCAGTCCAT	CTCTAATGGT	1804											
551	Gly	Ala	Asn	Gly	Thr	Ile	Thr	Leu	Gly	Ala	Leu	Ser	Thr	Leu	Thr	Glu
1805	TTCAATTAC	CGCTATCAGGGAAC	TTGGGGCCCAATCTTTGGG	GGACAGGGGCTTCTACG	1864											
571	Thr	His	Tyr	Gly	Tyr	Gln	Gly	Asn	Trp	---Gln	Leu	Ser	Trp	Ala	Asn	Ala
1865	---ACTGCA	CTTCACTGAGCTTAA	AACTGGCTATATTCCTA	ATCCGAGCGTATCGGC	1921											
590	Lys	Ile	Gly	Ser	Ile	Asn	Trp	Arg	Thr	Gly	Tyr	Ile	Pro	Ser	Pro	Glu
1922	TCTTTAGT	CCCTAATAGCTTATGA	ATGATTTAGATATAGCTCTCT	CCTCCATTATCTT	1981											
610	Asn	Leu	Pro	Leu	Asn	Ser	Leu	Trp	PGL	YAsn	Phe	Ile	Asp	Ile	Asn	

Qy		2042	AAC	TCTTCCATAAAGGATAGTACA	CAAAAACACGACGGGGTTTCGCCATTGTAGTGCGCGT	2100
Db		650	Asn	PhePheTyraArgaspSerMetProThr	ArghisgLyapheaRghisiliseSerGlyGly	669
Qy		2102	TAT	GTCATAGCAGGAACCTACATACT	TGTTTCAGATAAAGATTCTTAGTGCTGCATTTTGT	2161
Db		670	TyrAla	IeuGlyIleThraIathrThrPro	AlsagLueaspGlnIeuThrPheAlaPheCys	689
Qy		2162	CAG	CTCTTTTGGAAAGATAGAGACTACT	TTGTAGCTTAAGAATCAAGGTACAGTCTACGGA	2221
Db		690	GlnLeu	PheAlaArgaspArgAsnHisIle	ThrGlyLysAsnHisGlyAspThrTyrGly	709
Qy		2222	GG	AACTCTCTATTACAGCACACGAA	CAACCTATATCTCTCTTCCCTGC	2275
Db		710	AlaSer	LeuTyrPheHisIleThrGlu	GlyLeuPheAspIlealaAsnPheLeuTrpGly	739
Qy		2276	-----CGG	CTCTGTTCTGTTCTTATGTTC	TCTCCTACAGAGATTCCTGTTCTCTTT	2323
Db		730	LysAla	ThrArgAlaProTrpValLeu	SerGluIleSerGlnIleIleProLeuSerPhe	749
Qy		2324	TC	AGGAACCTTAGCTACACCACCAT	CGGATAACGATCGTAAAAACCAAGTATCAACATAT	2383
Db		750	AspAla	lysPheSerTyrLeuHisThr	AspAsnHieMetLysThrTyrThrAspAsn	769
Qy		2384	CCT	ACTGTTTAAAGAAAGCTGGG	GGAATGATAGTTCCTTTCCTTTCCTGCGGGAAGACGT	2443
Db		770	SerIle	IleLysGlySerTrpArgAsn	AspAlaPheCysalaspLeuGlyAlaSerLeu	789
Qy		2444	CCG	ATTGCTTAGTANAAGTGCTCTAT	TTTGACAGGTACATGCGCTTCATGAATTCGAG	2503
Db		790	ProPhe	VallileSerValProTyrLeu	LeuLysGluValGluProPheValLysValGln	809
Qy		2504	TTT	GTCATCCATCAGGAGGTTTTAA	AGAACACGGGAACAGAGCTCGTGAATTGGA	2563
Db		810	TyrIle	TyrAlahisGlnInasppTery	GluarGhisalagLuGlyArgalapeAsn	829
Qy		2564	AGT	AGCGCTTGTGTAATCTTGCTT	ACCTTACCTTCGGGATCCGATTTGATAGGAATCAGAC	2623
Db		830	LysSer	GluLeuIleasnValGluIle	ProIleGlyValThrPheGluIuaArgspSerLys	849
Qy		2624	TGC	CAAGATCGACAGTACATCTTA	CTCTTGGTTATACTCTGGATCTTGTCGTAGTAGTAAC	2683
Db		850	SerGly	luspGlyThrTyrAspLeuThr	LeuMetTyrIleLeuaspaLatYrArgArgAsn	869
Qy		2684	CCC	GACTGTACGACAAACACTGCG	AAATAGCGGTGATTCCTGGAAAAACCTTCGGTACGCAAT	2743
Db		870	ProLys	CysGlnThrSerLeulleAla	aSerAspAlaaAsnTrpMetAlaTyrGlyThrAsn	889
Qy		2744	TTG	CAAGACAGCTTTAGCTTCGT	CGGAGGNACCATTTTTCCTTTAACTCAAATTTT	2803
Db		890	LeuAla	argGlnGlyPheSerValArg	AlaAlaAsnHisPheGlnValAsnProHiMet	909
Qy		2804	GA	AGCCTTTAGCCAATTTCTTT	TGAATTCGGTGGGTCACTCTCGCAATTTACAATGTAGAC	2863
Db		910	GluIle	pheGlyGlnPheAlaPhe	GluValArgSerSerSerArgAsnTyrAsnThrAsn	929
Qy		2864	TT	AGGACAAAATACCAATTC	2884	
Db		930	LeuGly	SerLysPheCysPhe	2936	

```

RESULT 3
US-10-324-129-3
; Sequence 3, Application US/10324129
; Publication No. US20030157124A1
; GENERAL INFORMATION:
; APPLICANT: Murdin et al.
; TITLE OF INVENTION: Chlamydia antigens and corresponding DNA fragments and uses th
; FILE REFERENCE: 77813-4
; CURRENT APPLICATION NUMBER: US/10/324,129
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/113,439
; PRIOR FILING DATE: 1998-12-01

```

; PRIOR APPLICATION NUMBER: US 60/132,272
 ; PRIOR FILING DATE: 1999-05-03
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: Patent Ver. 2.0
 ; SEQ ID NO: 3
 ; LENGTH: 936
 ; TYPE: prt
 ; ORGANISM: Chlamydia pneumoniae
 ; US-10-324-129-3

Alignment Scores:		
Pred. No.:	6,38e-153	936
Score:	1965.00	411
Percent Similarity:	62.30%	Conservative: 179
Best Local Similarity:	43.40%	Mismatches: 327
Query Match:	36.33%	Indels: 30
DR:	12	Gaps: 16

US-09-428-122-1 (1-3000) x US-10-324-129-3 (1-936)

QY	101	ATGAAGTCTTCTTCCCAAGTTGTATTTTCTACATTTGCTATTTTC---CCTTTGTCT	157
Db	1	MetIySerSerValSerTriPLeuPhePheSerSerIleProLeuPheSerSerLeuSer	20
QY	158	ATGATTGCTACCGACACAGTTTTCGATTCAACT--GCGAGTTTCGATGGGAATAAAAAT	214
Db	21	IleValAlaAlaGluValThrLeuAspSerSerAsnAsnSerTyArgPcIySerAsnGly	40
QY	215	GGTAATTTTTCAGTTCGTGAGAGTCAGAA--GATGCTGGAACCTACCTACTTTTAAG	271
Db	41	ThrThrPheThrValPheSerThrThrAspAlaAlaAlaGlyThrThrTySerLeuLeu	60
QY	272	GGAATGTCACTCTAGAAAAATTTCTCGAAGACGACAGCAATCACAAAAGCTGTTTT	331
Db	61	SerAspValSerPheGlnAsnAlaGlyAlaLeuGlyIleProLeuAlaSerGlyCysPhe	80
QY	332	AACAACACTAAGGGCGATTTCGATTCTCAGGTAAACGGAACTCTATTGTTCCAAACG	391
Db	81	LeuGluAlaGlyGlyAspLeuThrPheGlnGlyAsnGlnHisAlaLeuLysPheAlaPhe	100
QY	392	GTGGATGACGAGGACTGTAGCAGGGGCTGCTGTTAACACGACGCGTGGTAGATAAAT	451
Db	101	IleAsnAlaGlySerSerAlaGlyThrValAlaSerThrSerAlaAlaAspLysAsnLeu	120
QY	452	ACGTTTATPAGGTTTTCTTCGCTACTCTTTATTTGCGTCTCTCGGAAGTTCGATACT	508
Db	121	LeuPheAsnAspPheSerArgLeuSerIleIleSerCysProSerLeuLeuLeuSerPro	140
QY	509	ACCGCAAGGAGCGGTAGCTGCTCTACGGGTAGCTTCAGTTTCACAAAAAATGTCACT	568
Db	141	ThrGlyGlnCysAlaLeuLys---SerValGlyAsnLeuSerLeuThrGlyAsnSerGln	159
QY	569	TTGCTCTTCAGCAAAACCTTTTCAACGGATATATGGCGGTGCTATCACCGCAAAAAC	628
Db	160	IleIlePheThrGlnAsnPheSerSerAspAsnGlyGlyValIleAsnThrLysAsnPhe	179
QY	629	TCATTAAACGGACACTACATGTCTGCTCTGTTTCTGAAATAAC-----TCCTCAAAG	682
Db	180	LeuLeuSerGlyThrSerGlnPheAlaSerPheSerArgAsnGlnAlaPheThrGlyLys	199
QY	683	AAAGCGGAGCCATTCCAGACTTCGATGCCCTTACCATTCTGGAACCAACGAGGGAAGTC	742
Db	200	GlnGlyGlyValValTyAlaThrGlyThrIleThrIleGluAsnSerProGlyIleVal	219
QY	743	TCTTTTTCGACAAATCTTCTTCGGATCTTCGGAGCTGCATATTTTTCACAGAGCCTCGGTG	802
Db	220	SerPheSerGlnAsnLeuAlaLysGlySerGlyGlyAlaLeuTySerThrAspAsnCys	239
QY	803	ACTATTCTTAATCTCTAAAGTTTCTTTATTGCAATAAGCTCACAGGAGGAGCTCC	862
Db	240	SerIleThrAspAsnPheGlnValIlePheAspGlyAsnSerAlaTrpGluAlaAlaGln	259
QY	863	TCAAACACGGGGGATGTGCAGGAGTGCTATCTGTCTTATAAACTAGTACAGATACT	922

Db	260	Ala-----GlnGlyAlaIleCysCys-----ThrThrThrAspLys 272
Qy	923	AAGGTCAACCTCACTCGAATCAGATGTTACTCTTCAGCAACAATACATCGACAACAGCG 982
Db	273	ThrValThrLeuThrGlyAsnLysAsnLeuSerPheThrAsnAsnThrAlaLeuThrTyr 292
Qy	983	GGAGGAGCTATCATCTGTAAGAGCTCGAATGGCTTCGGAGGAGCTACCCCTATTTCAGT 1042
Db	293	GlyGlyAlaIleSerGlyLeuLysValSerIleSerAlaGlyGlyProThrLeuPheGln 312
Qy	1043	AGAAATAGTGTCATCGAGGTACAGCTCTCAAA-----GGTGGAGCCATAGCTATCGAA 1096
Db	313	SerAsn---IleSerGlySerAlaGlyGlnGlyGlyAlaIleAsnIleAla 331
Qy	1097	GATAGTGGGAAATGAGTTATCCGCGATAGTGGTACATCTCTTTTATAGGAATACA 1156
Db	332	SerAlaGlyGlnLeuAlaLeuSerAlaThrSerGlyAspIleThrPheAsnAsnGln 351
Qy	1157	GTCACTTCTACTCTCTCGGAGCAATAGAAGTAGTAGTACGATAGGAACGAGTCAAAAG 1216
Db	352	ValThrAsnGlySerThrSerThr--ArgAsnAlaIleAsnIleIleAspThrAlaLys 370
Qy	1217	ATGACACCTTGGTCTCTGCTGCTGGTAGAGCCATCTACTTCTATGATCCATAACTACA 1276
Db	371	ValThrSerIleArgAlaAlaThrGlyGlnSerIleTyrPheTyrAspProIleThrAsn 390
Qy	1277	GGATCATCCACAACAGTTACAGATGCTTAAAGTTAATGAGCTCCGGCAGATTTCTGCA 1336
Db	391	ProGlyThrAlaAlaSerThrAspThrLeuAsnLeuAsnLeuAlaAspAlaAsnSerGlu 410
Qy	1337	CTACAATATACAGGAACATCATCTTCCAGGAGGAAAAAGTTATCAGAGACAGAGCCGCA 1396
Db	411	IleGluTyrGlyGlyAlaIleValPheSerGlyGluLysLeuSerProThrGluLysAla 430
Qy	1397	GATTTAAAAATCTTACTTCGAACTACTACAGCTCTAACTCTTTCAGGAGGTACTCTA 1456
Db	431	IleAlaAlaAsnValThrSerThrIleArgGlnProAlaValLeuAlaArgGlyAspLeu 450
Qy	1457	TCTTTAAACATGAGTGACTCTCGACACTCAGGCATTCTCACTCAACAGGCAGATTCTCGT 1516
Db	451	ValLeuArgAspGlyValThrValThrPheLysAspLeuThrGlnSerProGlySerArg 470
Qy	1517	CTCGAAATGGACGTAGGAATCTACTCTAGAACCTGCTGATACTAGC---ACCATAAACAAT 1573
Db	471	IleLeuMetAspGlyGlyThrLeuSerAlaLysGluAlaAsnLeuSerLeuAsnGly 490
Qy	1574	TTGGTCATTAACTCAGTTCTATAGACGGTCCAAAGAGGCAAAATAGAAACCAAAAGCT 1633
Db	491	LeuAlaValAsnLeuSerSerLeuAspGlyThrAsnLysAlaAlaLeuLysThrGluAla 510
Qy	1634	ACGTCAAAAATCTGACTTTATCTGGAAACCATCACTTTATTGGACCGACGGCAGCTTT 1693
Db	511	AlaAspLysAsnIleSerLeuSerGlyThrIleAlaLeuIleAspThrGluGlySerPhe 530
Qy	1694	TATGAAATCATAGTTTAAGAAATCCTCAGTCTCAGCATCTTAGAGCTCAAA----- 1747
Db	531	TyrGluAsnHisAsnLeuLysSerAlaSerThrTyrProLeuLeuGluLeuThrThrAla 550
Qy	1748	---GCTTCTGGAACCTGTAAACAGCACCCGACGTGATCTCCAGATCTCTTAATGGGTGAGAAA 1804
Db	551	GlyAlaAsnGlyThrIleThrLeuGlyAlaLeuSerThrLeuThrLeuGlnLeuProGlu 570
Qy	1805	TTCCATATGCGGTATCAGGAACTTGGGGCCCAATTGTTGGGGACAGAGGGCTTCTACG 1866
Db	571	ThrHisTyrGlyTyrGlnGlyAsnTrp---GlnLeuSerTrpAlaAsnAlaThrSerSer 589
Qy	1865	---ACTGCAACCTTCAACTGGAATAAACTGGCTATATTCTTAATCCCGAGCGATCCGCG 1921
Db	590	LysIleGlySerIleAsnTrpThrArgThrGlyTyrIleProSerProGluArgLysSer 609
Qy	1922	TCTTTAGTCCCTAATAGCTTATGGAATGCATTTATAGATATTTAGTCTCTCCATATCTT 1981

Db	610	AsnLeuProLeuAsnSerLeuTrpGlyAsnPheIleAspIleArgSerIleAsnGlnLeu	629
QY	1982	ATGAGAGACTGCAAAACGAAGGGTTTCAGAGGAGACCGTCTTTTGGTGTGCTGGATATATCT	2041
Db	630	IleGluThrLysSerSerGlyGluProPheGluArgGluLeuTrpLeuSerGlyIleAla	649
QY	2042	AAC TTC TCT CAT AAG GAT AGT ATC A A A A A C A G C G C G G T T T C C C A T T T G A G T G C G G T	2101
Db	650	AsnPhePheTyrArgAspSerMetProThrArgHisGlyPheArgHisIleSerGlyGly	669
QY	2102	T A N G T C A T A G G A G G A A C C T C A C A T A C T T G T T C A G A T A A G A T T C T T A G T G C T G A T T T T G T	2161
Db	670	TyrAlaLeuGlyIleThrAlaThrProAlaGluAspGlnLeuThrPheAlaPheCys	689
QY	2162	C A G C T C T T T G G A A G A G A T A G A G A C T A C T T T G T A G C T A A G A A T C A A G T A C A G T C T C A G G A	2221
Db	690	GlnLeuPheAlaArgAspArgAsnHisIleThrGlyLysAsnHisGlyAspThrTyrGly	709
QY	2222	G G A A C T C T C T A T T A C C A G C A C A G G A A A C C T A T A T C T C T C T C T T C C A A C T A	2275
Db	710	AlaSerLeuTyrPheHisHisThrGluGlyLeuPheAspIleAlaAsnPheLeuTrpGly	729
QY	2276	- - - - - C G G C C T T G T C C T C T C T T A T C T C T C A G A G A T T C C T G T T C T C T C T T	2323
Db	730	LysAlaThrArgAlaProTrpValLeuSerGluLeuSerGlnIleIleProLeuSerPhe	749
QY	2324	T C A G G A A A C C T T A G C T A C A C C C A T A C G G A T A C G A T C T G A A A C C A A G T A T A C A A C A T A T	2383
Db	750	AspAlaLysPheSerTyrLeuHisThrAspAsnHisMetLysThrTyrThrAspAsn	769
QY	2384	C C T A C T G T T A A G A A C T G G G G A A T A G A T T T C G C T T T A G A A T T C G G T G G A A G A C T	2443
Db	770	SerIleIleLysGlySerTrpArgAsnAspAlaPheCysAlaAspLeuGlyAlaSerLeu	789
QY	2444	C C G A T T T G C T T A G A T G A A G T C T A T T T G A G C A G T A C A T C G C C T T C A T G A A A T T C G A G	2503
Db	790	ProPheValIleSerValProTyrLeuLeuLysGluValGluProPheValLysValGln	809
QY	2504	T T T G T C T A T G C A C A T C A G G A A G T T T T A A G A A C A G G G A A C A G A A G T C G T G A A T T T G G A	2563
Db	810	TyrIleTyrAlaHisGlnAspPheTyrGluArgHisAlaGluGlyArgAlaPheAsn	829
QY	2564	A G T A G C G T C T T G T G A T C T T G C C T A C C G A T C C G G A T C C G A T T T G A T A G G A A T C A G A C	2623
Db	830	LysSerGluLeuIleAsnValGluIleProIleGlyValThrPheGluArgAspSerLys	849
QY	2624	T G C C A A G A T G C A A C T A A C T A A C T C T T G G T T A T A C T G T G G A T C T T G T C G T A G T A A C	2683
Db	850	SerGluLysGlyThrTyrAspLeuThrLeuMetTyrIleLeuAspAlaTyrArgArgAsn	869
QY	2684	C C G A C T G T A G C A A C A C T C G C A A T T A G C G G T G A T T C T T G G A A A A C C T T C G G T A C G A A T	2743
Db	870	ProLysCysGlnThrSerLeuIleAlaSerAspAlaAsnTrpMetAlaTyrGlyThrAsn	889
QY	2744	T T G C A C A G A C A G C T T T A G T C C T T G C G A G G A A C C A T T T T G C T T T A C T C A A A T T T	2803
Db	890	LeuAlaArgGlnGlyPheSerValArgAlaAlaAsnHisPheGlnValAsnProHisMet	909
QY	2804	G A A C C C T T T A G C C A A T T C T T T T G A T T G C G T G G T C A T C T C G C A A T T A C A A T G T A G A C	2863
Db	910	GluIlePheGlyGlnPheAlaPheGluValArgSerSerArgSerTyrAsnThrAsn	929
QY	2864	T T A G A G C A A A A T A C C A A T T C	2884
Db	930	LeuGlySerLysPheCysPhe	936

RESULT, T 4

US-09-452-380-4
 ; Sequence 4, Application US/09452380
 ; Patent No. US20020094340A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MURDIN, Andrew D.
 ; APPLICANT: COMEN, Raymond P.

; Sequence 4, Application US/10324129
; Publication No. US20030157124A1

; GENERAL INFORMATION:

; APPLICANT: Murdin et al.

; TITLE OF INVENTION: Chlamydia antigens and corresponding DNA fragments and uses there

; FILE REFERENCE: 77813-4

; CURRENT APPLICATION NUMBER: US/10/324,129

; CURRENT FILING DATE: 2002-12-20

; PRIOR APPLICATION NUMBER: US 60/113,439

; PRIOR FILING DATE: 1998-12-01

; PRIOR APPLICATION NUMBER: US 60/132,272

; PRIOR FILING DATE: 1999-05-03

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 4

; LENGTH: 925

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

; US-10-324-129-4

Alignment Scores:

Pred. No.:	2,12e-151	Length:	925
Score:	1946.50	Matches:	404
Percent Similarity:	62.54%	Conservative:	177
Best Local Similarity:	43.49%	Mismatches:	319
Query Match:	35.99%	Indels:	29
DB:	12	Gaps:	15

US-09-428-122-1 (1-3000) x US-10-324-129-4 (1-925)

```

Qy 152 TTGCTATGATGCTACCGAGAGACAGTTTGGATTCAAGT---GCGAGTTTCGATGGAAAT 208
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 8 LeuSerIleValAlaAlaGluValThrLeuAspSerSerAsnAsnSerThrAspGlySer 27
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 209 AAAAAATGCTAATTTTCAGTTCTGTGAGAGTCAGGAA---GATGCTCGAATACCTACCTA 265
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 28 AsnGlyThrThrPheThrValPheSerThrThrAspAlaAlaGlyThrThrTyros 47
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 266 TTTAAGGGAATGCTACTCTAGAAAATATTCTCGAACAGGACGACAGCAATCAAAAAGC 325
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 48 LeuLeuSerAspValSerPheGlnAsnAlaGlyAlaLeuGlyIleProLeuAlaSerGly 67
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 326 TGTATTAAACAACTAAGGCGGATTCAGTTTCTACAGGTAAAGCGGACTCTCTATTGTC 385
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 68 CysPheLeuGluAlaGlyAspLeuThrPheGlnGlyAsnGlnHisAlaLeuLysPhe 87
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 386 CAAACGGTGTGACGAGGACTGTAGCAGGCGCTGCTGTTAAACAGACGCGTGTAGATAAA 445
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 88 AlaPheIleAsnAlaGlySerSerAlaGlyThrValAlaSerThrSerAlaAlaAspLys 107
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 446 TCTACACGTTTATAGGTTTCTTCGCTATCTTTTATTCGCTCTCTCTCGAAGTTCGATA 505
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 108 AsnLeuLeuPheAsnAspPheSerArgLeuSerIleIleSerCysProSerLeuLeuLeu 127
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 506 ACT---ACCGGCAAGGAGCGGTAGTCTGTCTACGGGTAGCTTGAGTTTGACAAAAAT 562
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 128 SerProThrGlyGlnCysAlaLeuLys---SerValGlyAsnLeuSerLeuThrGlyAsn 146
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 563 GTCAGTTTGTCTTCACAAAACTTTTCAACGGATAATGCGGCTGCTATCCCGCAAAA 622
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 147 SerGlnIlePheThrGlnAsnPheSerSerAspAsnGlyGlyValIleAsnThrLys 166
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 623 ACTCTTTTCAATACAGGAGCTACAATGTCAGCTCTGTTTCTGAAATACC-----TCC 676
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 167 AsnPheLeuLeuSerGlyThrSerGlnPheAlaSerPheSerArgAsnGlnAlaPheThr 186
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 677 TCAAGAAAGCGGAGCGCATTCAGCTTCGATCGCCCTTACCATTAATCTGAAACAAAGG 736
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 187 GlyLysGlnGlyValValThrAlaThrGlyThrIleThrIleGluAsnSerProGly 206
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 737 GAAGTCTCTTTTCTGACAACTCTTCTCGGATTTCTGGAGCTGCAATTTTACAGAGCC 796
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 207 IleValSerPheSerGlnAsnLeuAlaLysGlySerGlyAlaLeuTyrosThrAsp 226
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Qy 797 TCGGTGACTATTCTAATATGCTAAAGTTTCTCTTATTGACAAATAGGTCACAGGAGCG 856
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 227 AsnCysSerIleThrAspAsnGlnValIlePheAspGlyAsnSerAlaTrpGluAla 246
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 857 AGTCTCTCAACAAACGGGGATATGTTCAGGAGGTGCTATCTGTGCTTTATAAACTAGTACA 916
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 247 AlaGlnAla-----GlnGlyGlyAlaIleCysCys-----ThrThrThr 259
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 917 GATACCTAAGGTCACCCCTCACTGAATCAGATGTTTACTCTTCAGCAACATACATCGACA 976
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 260 AspLysThrValThrLeuThrGlyAsnLysAsnLeuSerPheThrAsnAsnThrAlaLeu 279
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 977 ACAGCGGAGGAGCTATCTATGTAAGAAAGCTCGAATGCTCCGAGGAGCTTACCTTA 1036
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 280 ThrTyroGlyAlaIleSerGlyLeuLysValSerIleSerAlaGlyGlyProThrLeu 299
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 1037 TTCAGTAGAATAGTGTCAATGAGGTACAGCTCTCTAAA-----GGTGGAGCCATAGCT 1090
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 300 PheGlnSerAsn---IleSerGlySerSerAlaGlyGlnGlyGlyAlaIleAsn 318
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 1091 ATCGAAGATAGTGGGAATAGCTTTATCCGCGATAGTGGTGCACATTGCTTTTATAGGG 1150
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 319 IleAlaSerAlaGlyGluLeuAlaLeuSerAlaThrSerGlyAspIleThrPheAsnAsn 338
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 1151 AATACAGTCTCTTCTACTCTCTGGGACGAATAGAGTAGTAGTATGACATTAGCAACGAGT 1210
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 339 AsnGlnValThrAsnGlySerThrSerThr---ArgAsnAlaIleAsnIleAspThr 357
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 1211 GCAAGATGACAGCTTTCGCTTCTGCTGCTGAGAGCCATCTACTCTATGATCCCATTA 1270
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 358 AlaLysValThrSerIleArgAlaAlaThrGlyGlnSerIleThrPheTyrosProIle 377
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 1271 ACTCAGGATCATCCACACAGTTCAGATGCTCTTAAAGTTAATCAGACTCCGCGAGAT 1330
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 378 ThrAsnProGlyThrAlaAlaSerThrAspThrLeuAsnLeuAlaAspAlaAsn 397
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 1331 TCTCAGTACATATACAGGAGACATCTTCTCAGAGGAGAAAGTTATCAGAGACAGAG 1390
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 398 SerGluIleGluTyroGlyAlaIleValPheSerGlyGluLysLeuSerProThrGlu 417
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 1391 GCGCAGATCTTAAATCTTACTTCGAGCTACTCAGCTGTAACCTCTTTCAGGAGGT 1450
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 418 LysAlaIleAlaAlaAsnValThrSerThrIleArgGlnProAlaValLeuAlaArgGly 437
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 1451 ACTCTATCTTTAAACATCAGGTGCTCTCAGACTCAGGACTCTCAACAGGAGAT 1510
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 438 AspLeuValLeuArgAspGlyValThrValThrPheLysAspLeuThrGlnSerProGly 457
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 1511 TCTGCTCTCGAATGAGGAGTACCTACTCTAGAACCTCTGATAGCTAGC---ACATA 1567
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 458 SerArgIleLeuMetAspGlyGlyThrThrLeuSerAlaLysGluAlaAsnLeuSerLeu 477
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 1568 AACAAATTTGGTCAATTAACATCAGTTCTTATAGAGCGGTGCAAGGCAAAAATAGAAACC 1627
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 478 AsnGlyLeuAlaValAsnLeuSerSerLeuAspGlyThrAsnLysAlaAlaLeuLysThr 497
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 1628 AAAGCTAGTCAAAAATCTGACTTTATCTGGAACCATCTACTTATTTGACCCGAGCGGC 1687
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 498 GluAlaAlaAspLysAsnIleSerLeuSerGlyThrIleAlaLeuIleAspThrGluGly 517
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 1688 ACCTTTTATGAAATCATAGTTTAAAGAAATCCTCAGTCTTACGACATCTTAGAGCTCAAA 1747
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 518 SerPheTyroGluAsnHisAsnLeuLysSerAlaSerThrTyroLeuLeuGluLeuThr 537
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 1748 -----GCTCTGGAATGTAAACAGCACCGCAGTGCATCCAGATCTCTATAATGGGT 1798
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 538 ThrAlaGlyAlaAsnGlyThrIleThrLeuGlyAlaLeuSerThrLeuThrLeuGlnGlu 557
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 1799 GAGAAATCCATACGCTATCAGGGAACCTTGGGGCCCAATTGTTGGGGACAGCGGCT 1858
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 558 ProGluThrHisTyroGlyTyroGlnGlyAsnTrp---GlnLeuSerTrpAlaAlaThr 576
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

QY	1859	TCACG---ACTGCAACCTTCACTGGACCTAAACCTGGCTATATTTCTCTAATCCGAGCGT	1915
Db	577	SeqSerLysIleGlySerIleAsnTrpThrArgThrGlyTyrIleProSerProGluArg	596
QY	1916	ATCGGCTCTTGTAGTCCCTATAGCTATGTAATGCATTTATAGATATTTAGCTCTCTCCAT	1975
Db	597	LysSerAsnLeuProLeuAsnSerLeuTrpGlyAsnPhelleaspIleArgSerIleAsn	616
QY	1976	TATCTATGAGACTGCAACAGAGGTTGTCAGGAGACCGTGTCTTTTGGTGTGCTGGA	2035
Db	617	GlnLeuIleGluThrLysSerSerGlyGluProPheGluArgGluLeuTrpLeuSerGly	636
QY	2036	TTATCTAATCTTCTCCATAGGATAGTACAAAACAGACGCGGTTTCGCCATTTTCAGT	2095
Db	637	IleAlaAsnPhePheTyrArgAspSerMetProThrArgHisGlyPheArgHisIleSer	656
QY	2096	GGCGGTTATGTCATAGAGGAAACCTCATATCTTGTTCAGATAAGATTCTTAGTGTGCA	2155
Db	657	GlyGlyTyrAlaLeuGlyIleThrAlaThrTrpAlaGluAspGlnLeuThrPheAla	676
QY	2156	TTTTGTCTAGCTCTTTGGAAGAGATAGAGACTPACTTTGTAGCTAAGAAATCAAGGTACAGTC	2215
Db	677	PheCysGlnLeuPheAlaArgAspArgAsnHisIleThrGlyLysAsnHisGlyAspThr	696
QY	2216	TACGAGGAATCTCTATTACGACGACGACAAACCTATATCTCTCTCTCTGCAAACTA	2275
Db	697	TyrGlyAlaSerLeuTyrPheHisHisThrGluGlyLeuPheAspIleAlaAsnPheLeu	716
QY	2276	-----CGCGCTTGTCTGTCTTATGTTCTCTACAGAGATTCTCTGTT	2317
Db	717	TrpGlyLysAlaThrArgAlaProTrpValLeuSerGluIleSerGlnIleProLeu	736
QY	2318	CTCTTTTTCAGGAAACCTTAGCTACACCATACGGATAACGATCTGAAACCAAGTATACA	2377
Db	737	SerPheAspAlaLysPheSerTyrLeuHisThrAspAsnHisMetLysThrTyrTyrThr	756
QY	2378	ACATATCTCTGTTAAAGAACTGGGGGATGATAGTTTCGCTTTTGAATTCGGTGA	2437
Db	757	AspAsnSerIleIleLysGlySerTrpArgAsnAspAlaPheCysAlaAspLeuGlyAla	776
QY	2438	AGAGCTCCGATTTCGTAGTAAAGTGTCTATTGAGCAGTACATGCCCTTCATGAAA	2497
Db	777	SerLeuProPheValIleSerValProTyrLeuLeuLysGluValGluProPheValLys	796
QY	2498	TTGAGTTTGTCTATGCATCAGGAAGTTTAAAGAACAGGGAACAGAGCTCGTGAA	2557
Db	797	ValGlnTyrIleTyrAlaHisGlnGlnAspPheTyrGluArgHisAlaGluGlyArgAla	816
QY	2558	TTTGAAGTACGGCTTGTGATCTTGCTTACCTATCGGATCCGATTTCGATAAGGAA	2617
Db	817	PheAsnLysSerGluLeuIleAsnValGluIleProIleGlyValThrPheGluArgAsp	836
QY	2618	TCAGACTGCCAGATGCAACGCTACATCTAATCTTGTGTTATATCTGTGATCTTGTCT	2677
Db	837	SerLysSerGluLysGlyThrTyrAspLeuThrLeuMetTyrIleLeuAspAlaTyrArg	856
QY	2678	AGTAACCCGACTGTACGACAACTCGCAATTTAGCGGTGATTTTGGAAACCTTCGCT	2737
Db	857	ArgAsnProLysCysGlnThrSerLeuIleAlaSerAspAlaAsnTrpMetAlaTyrGly	876
QY	2738	ACGAATTTGGCAAGACAGCTTTAGTCTTCGTGAGGAGCAACATTTTTCCTTACTCA	2797
Db	877	ThrAsnLeuAlaArgGlnGlyPheSerValArgAlaAlaAsnHisPheGlnValAsnPro	896
QY	2798	AATTTTGAACCTTTAGCCAAATTTCTTTTGAATTCGGTGGGTCTATCTCGCAATTTACAAT	2857
Db	897	HisMetGluIlePheGlyGlnPheAlaPheGluValArgSerSerArgAsnTyrAsn	916
QY	2858	GTAGACTTAGGAGCAAAATACCAATTC	2884
Db	917	ThrAsnLeuGlySerLysPheCysPhe	925
QY	671	ACCTCTCAAGAAAGGCGGAGCCATTTCAGACTTCGATCCCTTACCATTACTGGAAC	730

Db	190	SerThrAlaLysLysGlyGlyAlaIleAlaAlaThrGlyIleAlaHisLeuSerAsp	209
Qy	731	CARGGGGAAGTCTCTTTTCTGCAATACTCTTCGGATTCTGGAGCTGCAATTTTACA	790
Db	210	GlnGlyThrIleArgPheSerGlyAsnThrAlaValAsnSerGlyGlyAlaValTyr	229
Qy	791	GAAGCCTCGGTAGTATTCTTAATAATGCTAAAGTTTCTCTTATTGACAAATAGGTACA	850
Db	230	GluAlaSerMetThrIleAlaGlyAsnAsnHisAlaAlaPheSerAsnAsnAlaValSer	249
Qy	851	GGAGCGAGCTCTCAACAACGGGGGATATGTCAGGAGGTGCTATCTGTCTTATAAACT	910
Db	250	GlySerSer-----AspGlyCysGlyAlaIleHisCysSerLysThr	264
Qy	911	AGTACGATACTAAGGTACCCCTCACTCGAAATCAGATGTTACTCTTCAGCAACATACA	970
Db	265	GlySerAlaProThrLeuThrIleArgAspAsnLysValLeuIlePheGluGluAsnThr	284
Qy	971	TCGACACACGGGGAGGAGTATCTATCTGAAAGAAGCTCGAAGTCGGCTTCGGAGACTT	1030
Db	285	SerSerAlaLysGlyGlyAlaIleTyrThrAspLysLeuIleLeuThrSerGlyGlyPro	304
Qy	1031	ACCCATTCTAGTAGAATAGTGTCAATGGAGGTACAGCTCCTAAAGTGGAGCCATAGCT	1090
Db	305	ThrAlaPheIleAsnAsnLysValThrHisAlaThr---ProLysGlyGlyAlaIleGly	323
Qy	1091	ATCGAAGATAGTGGGAATTGAGTTTATCCGCGCATAGTGGTGACATGTGCTTTTAGGG	1150
Db	324	IleAlaAlaAsnGlyGlyCysSerLeuThrAlaGluHisGlyAspIleThrPheAspAsn	343
Qy	1151	AATACAGTCACTTCTACTCTCTCTGGGAGC--AATAGTAGTAGTAGTTCAGATTAGGAACG	1207
Db	344	AsnLeuMetAlaThrGlnAspAsnAlaThrIleLysArgAsnAlaIleAsnIleGluGly	363
Qy	1208	AGTGCAAGATGACAGCTTTGCGTTCGTGCTGGTAGACCCATCTACTTCTATGATCCC	1267
Db	364	AsnGlyLysPheValAsnLeuArgAlaAlaSerGlyLysThrIleSerPheTyrAspPro	383
Qy	1268	ATACTACAGGATCATCCACAACAGTTTACAGATCTCTTAAAGTTAATGAGACTCCGGCA	1327
Db	384	IleThrValGluGlyAsn-----AlaAlaAspLeuLeuThrLeuAsnLysAlaGluGly	401
Qy	1328	GATTCTGCACATAATACAGGAAACATCATCTTCACAGGAGAAAGTTATCAGAGACA	1387
Db	402	AspLysThr-----TyrAsnGlyArgIleIlePheSerGlyGlyLysLeuThrGluGlu	419
Qy	1388	GAGGCCGAGATCTAAATATCTTACTTCGAAGCTACTACAGCTGTAACTCTTTCAGGA	1447
Db	420	GlnAlaAlaValAlaAspAsnLeuLysThrThrPheThrGlnProIleThrLeuAlaAla	439
Qy	1448	GGTACTCTATCTTTAAACATGAGTGACTCTGCAGACTCAGGCATTTCATCAACAGGCA	1507
Db	440	GlyGluLeuValLeuArgSerGlyValGluValGluAlaLysThrValGlnThrAla	459
Qy	1508	GATTCTCGTCTCGAAATGGAGTAGGAACTACTCTAGAA---CCTGCTGATACTAGCACC	1564
Db	460	GlySerLeuIleLeuMetAspAlaGlyThrLysLeuSerAlaLysThrGluAspAlaThr	479
Qy	1565	ATAACCAATTGGTTCATTAACATCAGTTCTTATAGACGGTCAAGAGGCAAAATAGAA	1624
Db	480	LeuThrAsnLeuAlaIleAsnProAsnThrLeuAspGlyLysLysPheAlaValValAsp	499
Qy	1625	ACCAAAGTTCGTCAAAATATCTGACTTTTATCTGGAAACCATTCATTTATTGGACCGAGC	1684
Db	500	AlaValAlaAlaGlyLysAsnValThrLeuSerGlyAlaIleGlyValIleAspProThr	519
Qy	1685	GGCAGGTTTATGAATCATAGTTTAAAGAAATCCTCAGTCCCTACGACATCTTAGAGCTC	1744
Db	520	GlyLysPheTyrGluAsnHisLysLeuAsnAspThrLeuAlaLeuGlyGlyIleGlnLeu	539
Qy	1745	AAAGCTTCTGGAATGTAAACAGCAGCGAGTCCAGATCTCCATATCTAATATGGGT---GAG	1801

Db	540	SerGlyLysGlySerValThrThrAsnVal	--ProSerHisValValGlyValAla	558
QY	1802	AAATTCCATTAGCGCTATCAGGAACTTGGGCCCAATGTTGG	-----	1846
Db	559	GluThrHisTyrGlyTyrGlnGlyAsnTrpSer	--ValSerTrpValLysAspAsnAsn	577
QY	1847	GGACAGGGCTTCTACGACTGCAACTTCACCTGAGCTAAAACTGGCTATATTCCTAAT	1906	
Db	578	SerAspProLysThrGlnThrAlaIlePheThrTrpAsnLysThrGlyTyrValProAsn	597	
QY	1907	CCGACGCTATCGGCTCTTATGTCCTTAATAGCTTATGGAATGCATTTATAGATATATAGC	1966	
Db	598	ProGluArgAlaProLeuValLeuAsnSerLeuTrpGlySerPheIleAspLeuArg	617	
QY	1967	TCCTCCATATCTTATGAGACTGCAACAGAAAGG	---TTGCAGGAGGACCGTGCTTTT	2023
Db	618	SerIleGlnAspValLeuGluArgSerValAspSerIleLeuGlnThrArgArgGlyLeu	637	
QY	2024	TGCTGTGCTGGATTACTAACTTCTTCATAAGGATAGTACAAAAACAGACGCGGTTT	2083	
Db	638	TrpValSerGlyIleGlyAsnPhePheHisLysAspArgAsnAlaGluAsnArgLysPhe	657	
QY	2084	CGCATTTAGTGGCGGTATGTATGATGAGGAGAACTATACATCTTGTTCAGATAAGATT	2143	
Db	658	ArgHisIleSerSerGlyTyrValLeuGlyAlaThrThrAsnThrSerArgGluAspSer	677	
QY	2144	CTTAGTCCTGCATTTGTCTCAGCTCTTTGGAGAGATAGAGACTACTTTGTAGCTAGAAT	2203	
Db	678	LeuSerValAlaPheCysGlnLeuPheAlaLysAspLysAspTyrLeuValSerLysAsn	697	
QY	2204	CAAGGTACAGCTACGAGGAACTCTCTATTACCAGCACAAACCACTATATCTCTCTT	2263	
Db	698	AlaAlaAsnValTyrAlaGlySerValTyrTyrGlnHisValSerLysPheAspAspLeu	717	
QY	2264	-----CCTTGCAAACTAGGCGCTGTTCGTTCCTTATGTT	2299	
Db	718	ThrArgLeuPheAsnGlyProAsnThrCys	-----CysSerGlyPhe	731
QY	2300	CCTACAGAGATTCCTGTCTCTTTTCAGGAAACCTTAGCTACACCCATACGGATAACGAT	2359	
Db	732	SerLysGluIleProllePheLeuAspAlaGlnIleThrTyrCysHisThrAlaAsnAsn	751	
QY	2360	CTGAACCAAGATATACACATATCTACTGTGTTAAAGGAAGCTGGGGAATGATAGTTTC	2419	
Db	752	MetThrThrSerTyrThrAspTyrProGluValLysGlySerTrpGlyAsnAspThrLeu	771	
QY	2420	GCCTTAGAATTCGTTGGAGAGCTCGGATTGCTGTAGATCAAAAGTGCCTATTGTAGCAG	2479	
Db	772	GlyLeuThrLeuSerThrSerValProlleProValPheSerSerSerIlePheAspSer	791	
QY	2480	TACATGCCCTTCAGAAATTGCGATTTGCTATGTCACATCAGGAAGGTTTAAAGAACAG	2539	
Db	792	TyrAlaProPheAlaLysLeuGlnValTyrAlaHisGlnAspAspPheLysGluPro	811	
QY	2540	GGAACGAAGCTCGTGAATTTGGAAGTAGCGGCTTGTGTAATCTTGCTTACCTATCGGG	2599	
Db	812	ThrThrGluGlyArgValPheGluSerSerAspLeuLeuAsnValSerValProlleGly	831	
QY	2600	ATCCGATTTGATAGGAATCAGACTGCCAGATGCAACGTACAACTTCAACTCTTTGGTTAT	2659	
Db	832	IleLysPheGluLysLeuSerTyrGlyGluArgSerAlaTyrAspLeuThrLeuMetTyr	851	
QY	2660	ACTGTGATCTGTTCGTAGTAACCCGACTGTACGACACACTGCGAATTACGGGTGAT	2719	
Db	852	IleProAspValTyrArgHisAsnProSerCysMetThrGlyLeuAlaIleAsnAspVal	871	
QY	2720	TCITGGAAAACCTTCGGTAGCAATTTGGCAAGACAAAGCTTTAGTCCCTTCGTGAGGGAAC	2779	
Db	872	SerTrpLeuThrAlaThrAsnLeuAlaArgGlnAlaPheIleValArgAlaGlyAsn	891	
QY	2780	CATTTTGTCTTAACTCAAAATTTTGAAGCCTTTAGCCAAATTTTCTTTGAAATTCGCTGGG	2839	
Db	892	HisIleAlaLeuThrSerGlyValGluMetPheSerGlnPheGlyPheGluLeuArgSer	911	


```

QY 2840 TCATCTCGCAATTAACAAGTAGACTTAGGAGCAAAAATACCAATTC 2884
Db |||||
QY 912 SerSerArgAsnTyrAsnValAspLeuGlyAlaLysValAlaPhe 926
Db |||||

RESULT 7
US-10-023-437-57
; Sequence 57, Application US/10023437
; Publication No. US20020183272A1
; GENERAL INFORMATION:
; APPLICANT: JOHNSTON, STEPHEN A.
; APPLICANT: STEWKE-HALE, KATHERINE
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: KALTENBOECK, BERNHARD
; TITLE OF INVENTION: METHODS AND COMPOSITIONS for Vaccination COMPRISING NUCLEIC ACID
; TITLE OF INVENTION: AND/OR POLYPEPTIDE SEQUENCES OF CHLAMYDIA
; FILE REFERENCE: US/10/023,437
; CURRENT APPLICATION NUMBER: US/10/023,437
; PRIORITY FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/225,839
; PRIORITY FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
; LENGTH: 926
; TYPE: PRT
; ORGANISM: Chlamydia psittaci
US-10-023-437-57

Alignment Scores:
Pred. No.: 8,3e-149 Length: 926
Score: 1915.00 Matches: 410
Percent Similarity: 59.27% Conservative: 156
Best Local Similarity: 42.93% Mismatches: 333
Query Match: 35.40% Indels: 56
Db: 14 Gaps: 18

US-09-428-122-1 (1-3000) x US-10-023-437-57 (1-926)
QY 101 ATGAAGTCTCTTCCCAAGTTTGTTATTTCTACATTTGCTATTTCCCTTTGCTATG 160
Db |||||
QY 1 MetArgProSerLeuTyrLysIleLeuIleSerSerThrLeuThrLeuProIleSerPhe 20
QY 161 -----ATTGTCACGACACAGTTTGGATTCAAGT 190
Db |||||
QY 21 HisPheSerGlnLeuHisAlaGluValAlaLeuThrGlnGluSerIleLeuAspAla--- 39
QY 191 GCGAGTTTCGATGGGAATAAATGTTATTTTCAGTTTCGTGAGAGTCAGGAGATGCT 250
Db |||||
QY 40 -----AsnGlyAlaPheSerProGlnSerThrSerThrAlaGly 52
QY 251 GGAACCTACCTACTATTTAAGGGAAATGTCACTCTAGAAAAATATTCCTGGAAACAGGCACA 310
Db |||||
QY 53 GlyThrIleTyrAsnValGluSerAspIleSerIleValAspVal---GlyGlnThrAla 71
QY 311 GCAATCACAAAAGCTGTTTAAACAACATAAGGGCGATTGCTTTCAGGTACAGTACGGG 370
Db |||||
QY 72 AlaLeuAlaSerSerAlaPheValGlnThrAlaAspAsnLeuThrPheLysGlyAsnAsn 91
QY 371 AACTCTCTATTGTTCCAAACGGTGGATGAGGAGCTGTAGCAGGGCTGCTGTTAACAGC 430
Db |||||
QY 92 HisSerLeuSerIleThrAsnAlaAsnAlaGly---AlaAsnProAlaGlyIleAsnVal 110
QY 431 AGCGTGTAGATAAATCTACCAAGTTTATAGGGTTTCTTCGCTATCTTTTATTGCGTCT 490
Db |||||
QY 111 AsnThrAlaAspLysIleLeuThrLeuThrAspPheSerLysLeuSerPheLysGluCys 130
QY 491 CTTGGAAAGTTTCGATAACTACCGGCAAGAGCGGTAGCTGCTCTACGGGTAGCTTGAGT 550
Db |||||
QY 131 ProSerSerLeuValAsnThrGlyLysGlyAlaMetLys---SerGlyGlyAlaLeuAsn 149
QY 551 TTGACAAAAAATGTCAGTTTCTCTTCAGCAAAAACCTTTTCAACGGGATAATGCGGTGCT 610
Db |||||

```

```

150 LeuAlaAsnAsnAlaSerIleLeuPheAspGlnAsnTyrSerAlaGluAsnGlyGlyAla 169
611 ATCACCGCAAAAACCTCTTTTCATTAACAGGAGCTACCAATGTGAGCTCTGTCTTCTGAAAAT 670
Db |||||
170 IleSerCysLysAlaPheSerLeuThrGlySerSerLysGluIleSerPheThrThrAsn 189
QY |||||
671 ACCTCTCTCAAGAAAGAGCGGAGCCATTACAGACTTCGATGCCCTTACCATTACTTGGAAAC 730
Db |||||
190 SerThrAlaLysLysGlyGlyAlaIleAlaAlaThrGlyIleAlaHisLeuSerAspAsn 209
QY |||||
731 CAAGGGGAAGTCTCTTTTCTGACATACTCTTCGGAATCTCGAGCTCTGGAGCTGCAATTTTACA 790
Db |||||
210 GlnGlyThrIleArgPheSerGlyAsnThrAlaValAsnSerGlyGlyAlaValTyrSer 229
QY |||||
791 GAAGCCCTCGGTGACTATTCTTAATAATGCTAAAGTTTCCTTTATTGACATAAGGTGACACA 850
Db |||||
230 GluAlaSerMetThrIleAlaGlyAsnAsnHisValAlaPheSerAsnAsnAlaValSer 249
QY |||||
851 GGAGCGAGCTCTCTCAACACAGGGGGATATGTTCAGAGGTGCTATCTGTGCTTTATAAACT 910
Db |||||
250 GlySerSer-----AspGlyCysGlyGlyAlaIleHisCysSerLysThr 264
QY |||||
911 AGTACAGATACTAAGGTCAACCTCAGTGGAAATCAGATGTTACTCTTCAGCAACAATACA 970
Db |||||
265 GlySerAlaProThrLeuThrIleArgAspAsnLysValLeuIlePheGluGluAsnThr 284
QY |||||
971 TCGACAAACAGCGGAGGAGCTATCTATGTGAAAAGCTCGAACTGGCTTCGAGGAGCTT 1030
Db |||||
285 SerSerAlaLysGlyGlyAlaIleTyrThrAspLysLeuIleLeuThrSerGlyGlyPro 304
QY |||||
1031 ACCCTATTTCAGTAGAAAATAGTGTCAATGGAGGTACAGCTCCTAAAGTGGAGCCATAGCT 1090
Db |||||
305 ThrAlaPheIleAsnAsnLysValThrHisAlaThr---ProLysGlyGlyAlaIleGly 323
QY |||||
1091 ATCGAAGATAGTGGGAATGAGTTTATCCCGCATGATGCTGATGCTTCTTTTAGGG 1150
Db |||||
324 IleAlaAlaAsnGlyGluCysSerLeuThrAlaGluHisGlyAspIleThrPheAspAsn 343
QY |||||
1151 AATCAGTCACTCTTACTACTCTCTGGGAGC---AATAGAAGTAGTATCGACTTAGGAACG 1207
Db |||||
344 AsnLeuMetAlaThrGlnAspAsnAlaThrIleLysArgAsnAlaIleAsnIleGluGly 363
QY |||||
1208 AGTCAAAAGATGACACAGCTTTGCTGCTGCTGTAGAGCCATCTACTTCTTATGATCCC 1267
Db |||||
364 AsnGlyLysPheValAsnLeuArgAlaAlaSerGlyLysThrIleSerPheTyrAspPro 383
QY |||||
1268 ATAACTACAGGATCATCCAAACAGTGTACAGTGTCTTAAAGTTAATCAGACTCCGGCA 1327
Db |||||
384 IleThrValGluGlyAsn-----AlaAlaAspLeuLeuThrLeuAsnLysAlaGluGly 401
QY |||||
1328 GATTCCTGCATACATAATACAGGGAACATCATCTTCACAGGAGAAAAGTTATCAGAGACA 1387
Db |||||
402 AspLysThr-----TyrAsnGlyArgIleIlePheSerGlyLysLeuThrGluGlu 419
QY |||||
1388 GAGGCGCGCAGATTCTTAAATACTTACTTCTGAAAGCTACTACAGCTGTAACTCTTTCAGA 1447
Db |||||
420 GlnAlaAlaValAlaAspAsnLeuLysThrPheThrGlnProIleThrLeuAlaAla 439
QY |||||
1448 GGTACTCTTACTTTAAACATCGAGTGTCTGCGAGACTCAGCGCATCTCACTCAACAGGCA 1507
Db |||||
440 GlyGluLeuValLeuArgSerGlyValGluValGluAlaLysThrValValGlnThrAla 459
QY |||||
1508 GATTCCTCGTCTCGAATGAGCTAGGAACTACTCTAGAA---CCTGCTGATACTAGCAC 1564
Db |||||
460 GlySerLeuIleLeuMetAspAlaGlyThrLysLeuSerAlaLysThrGluAspAlaThr 479
QY |||||
1565 ATAAACAAATTTGGTCAATTAACATCATGTTCTATAGAGCGGTGCAAGGCAAAAATAGAA 1624
Db |||||
480 LeuThrAsnLeuAlaIleAsnProAsnThrLeuAspGlyLysPheAlaValValAsp 499
QY |||||
1625 ACCAAAGCTACGTCAAAAATCTGACTTTTATCTGGAACCATCTATCTTATGACCCGAGC 1684
Db |||||
500 AlaValAlaAlaGlyLysAsnValThrLeuSerGlyAlaIleGlyValIleAspProThr 519

```


389	IleThrLeuThrAsnLeuAspIleasnIleAlaSerLeuGlyGlyGlyThrSer	408
1610	AAGCGAAAAATAGAAACCAAGACTACGTCAAAAATCTGACTTTATCTGGAACCATCACT	1669
409	ProAlaLysLeuAlaThrAsnThrAlaSerGlnAlaIleThrIleAsn---AlaValAsn	427
1670	TTATTGGACCGGACGGGACGCTTTTATCAAATCATAGTTTAAAGAAATCCCTCAGTCCTAC	1729
428	LeuValAspAlaAspGlyAsnAlaTyrGluAspProIleLeuAlaThrSerLysProPhe	447
1730	GACATCTTAGAGCTCAAAAGCTTCTGGAACTGTAAACAAGCAGCGCACTCAGATCCATCCT	1789
448	ThrAlaIle-----ValAlaThrThrAsnAlaSerThrValThrGlnProThrAsp	464
1790	ATAATGGGTGAG-----AAATTCAATTACGGCTATCAGGGAACCTTGGGCCCAATT	1840
465	AsnLeuThrAsnTyrValProProThrHisTyrGlyTyrGlnGlyAsnTrp---ThrVal	483
1841	GTTTGGGGACAGGGGCTTCTACG---ACTGCAACTCTCAACTGCACATAAAATCGCTAT	1899
484	ThrTrpAspThrGluThrAlaThrLysThrAlaThrLeuThrTrpGluGlnThrGlyTyr	503
1898	AATCTCTAAATCCCGAGCGATCCGGCTCTTTAGTCCCTAATAGCTTTATGGAATCAATTATA	1957
504	SerProAsnProGluArgGlnGlyProLeuValProAsnThrLeuTrpGlyAlaPheSer	523
1958	GATATTAGCTCTCTCCATTATCTTATGAGAGACTGCMAAGAAAGGTTGCAGGAGACCGT	2017
524	AspLeuArgAlaIleGlnAsnLeuMetAspIleSerValAsnGlyAlaAspTyrHisArg	543
2018	GCTTTTGGTGTGCTGGATTACTTAACCTCTTCCATAAGATAGTACAAAACACAGACCG	2077
544	GlyPheTrpValSerGlyLeuAlaAsnPheLeuHisLysSerGlySerAspThrLysArg	563
2078	GGGTTTCGGCATTTAGTGGCGGTATGTATAGAGAGAGAAACCTACATCTTGTTCAGAT	2137
564	LysPheArgHisAsnSerAlaGlyTyrAlaLeuGlyValTyrAlaLysThrProSerAsp	583
2138	AAGATCTTTAGTCTGCATTTTGTCTAGCTCTTTGGGAAGATAGACACTCTTTGTAGCT	2197
584	AspIlePheSerAlaAlaPheCysGlnLeuPheGlyLysAspLysAspTyrLeuValSer	603
2198	AGAATCAAGGTACACTACGGAGAACTCTCTATTACGACACACGAAACCTATATC	2257
604	LysAsnAlaAlaAsnIleTyrAlaGlySerLeuTyrTyrGlnHis-----Ile	619
2258	TCCTCTTCCTTGCAAACTACGGCCCTTGTCTGTCTTATGTTCTCCTACAGAGATTCCTGTT	2317
620	SerTyrTrpSerAlaTrpGlnAsnLeuLeuGlnAsnThrIleGlyAlaGluAlaProLeu	639
2318	CTCTTTTCAGGAAACCTTAGCTACACCCATACGGATACGATCTGAAACCC-----	2368
640	ValLeuAsnAlaGlnLeuThrTyrCysHisAlaSerAsnAspMetLysThrAsnMetThr	659
2369	-----AAGTATACAACATATCTCTCTGTTTAAAGAGAGCTGGGGGATGAT	2413
660	ThrThrTyrAlaProArgLysThrThrTyrAlaGluIleLysGlyAspTrpGlyAsnAsp	679
2414	AGTTTCGCTTTAGAAATTCGGTGGAGAGCTCCGATTTGCTTAGATGAAGTCTCTATTT	2473
680	CysPheGlyValGluLeuGlyAlaThrValProIleGlnThrGluSerSerLeuLeuPhe	699
2474	GAGCAGTACATGCCCTTCATGAAATTCGAGTTTGTCTATGCACATCAGGAAGGTTTAAA	2533
700	AspMetTyrSerProPheLeuLysPheGlnLeuValHisThrHisGlnAspPheLys	719
2534	GAACAGGGGAACA---GAAGCTCGTGAATTTGGAGATGACCGCTCTTGTAATCTTCGCTTA	2590
720	GluAsnAsnSerAspGlnGlyArgTyrPheGluSerSerAsnLeuThrAsnLeuSerLeu	739
2591	CCTATCCGGATCCGATTGTATAGGAATCAGACTGCCCAAGATGCAACGTAACAATCTAACT	2650

Db 740 ProIleGlyLeuIysPheGluArgPheAlaAsnAspThrAlaSerTyrHisValThr 759
 QY 2651 CTTGGTATACGTGGATCTTTCGTAGTAACCCCGACTGTACGACAACTCGGAATT 2710
 Db 760 AlaAlaTyrSerProAspValArgSerAsnProAspCysThrThrSerLeuLeuVal 779
 QY 2711 AGCGGTGATTCT-----TGGAAACCTTCGGTAGCAATTTGGCAAGCAAGCTTTAGTC 2764
 Db 780 SerProAspSerAlaValTyrValThrLysAlaAsnAsnLeuAlaArgSerAlaPheMet 799
 QY 2765 CTTCTGCGAGGAACCAATTTTTCCTTAACTCAAAATTTTGAAGCCTTTAGCAATTTCT 2824
 Db 800 LeuGlnAlaGlyAsnTyrLeuSerLeuSerHisAsnIleGluIlePheSerGlnPheGly 819
 QY 2825 TTTGAATTCGGTGGTCATCTCGCAATTAACAATGTAGACTTAGGAGCAAAATACCAATTC 2884
 Db 820 PheGluLeuArgGlySerSerArgThrTyrAsnValAspLeuGlySerLysIleGlnPhe 839

RESULT 9

US-10-023-437-23
 ; Sequence 23, Application US/10023437
 ; Publication No. US20020183272A1
 ; GENERAL INFORMATION:
 ; APPLICANT: JOHNSTON, STEPHEN A.
 ; APPLICANT: STEMKE-HALE, KATHERINE
 ; APPLICANT: SYKES, KATHRYN F.
 ; APPLICANT: KALTENBOECK, BERNHARD
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING NUCLEIC ACID
 ; FILE REFERENCE: AND/OR POLYPEPTIDE SEQUENCES OF CHLAMYDIA
 ; FILE REFERENCE: US/10/023,437
 ; CURRENT APPLICATION NUMBER: US/10/023,437
 ; CURRENT FILING DATE: 2001-12-17
 ; PRIOR APPLICATION NUMBER: 60/225,839
 ; PRIOR FILING DATE: 2000-12-15
 ; NUMBER OF SEQ ID NOS: 69
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 23
 ; LENGTH: 839
 ; TYPE: PRT
 ; ORGANISM: Chlamydia psittaci
 US-10-023-437-23

Alignment Scores:
 Pred. No.: 8,64e-128 Length: 839
 Score: 1659.50 Matches: 384
 Percent Similarity: 54.17% Conservatives: 136
 Best Local Similarity: 40.00% Mismatches: 287
 Query Match: 30.68% Indels: 153
 DB: 14 Gaps: 25

US-09-428-122-1 (1-3000) x US-10-023-437-23 (1-839)

QY 89 ATATATTTTACATGAAGTCTTCTTCCCAAGTTGTATTTCATATTGCTATTTC 148
 Db 5 ValTyrTrpPheLeuIleSerSer-----SerLeuPheAlaSerAsn 18
 QY 149 CTTTGTCTATGATT---GCTACCGACACAGTTTTCGATTCAGTCCGAGTTTCGATGG 205
 Db 19 SerLeuSerPheAlaAsnAspAlaGlnThrAlaLeuThrProSerAspSerTyrAsnGly 38
 QY 206 AAT---AAAAATGGTAATTTTTCAGTTTCGTGAGAGTCAGGAAGATCTGGAACCTACTAC 262
 Db 39 AsnValThrSerGluGluPheGlnValLysGluThr-----SerSerGlyThrThrTyr 56
 QY 263 CTATTTAAGGGAATGTCACTCTAGAAAATATTCCTGGACAGGACAGCAATCAAAA 322
 Db 57 ThrCysGluGlyAsnValCysIle---SerPheAlaGlyLysAspSerGlyLeuLysLys 75
 QY 323 AGCTGTTTTAACACACTAAGGGCGATTTCGACTTTTCACAGGTAAACGGGAACCTCTCTATTG 382
 Db 76 SerCysPheSerAlaThr---AspAsnLeuThrPheLeuGlyAsnGlyTyrThrLeuCys 94
 QY 383 TTCAAACGGTGGATGACAGGACTGTAGCAGGGGCTGCTGTTAACACAGCGGTGATAT 442

Db 95 PheAspAsnIleThrThrThrAlaSerAsnProGlyAlaIleAsnValGlnGlyGlnGly 114
 QY 443 AAATCTACACAGTTTATAGGGTTTCTCGCTATCTTTTATTCGCTCTCCTCGAAGTTTCG 502
 Db 115 LysThrLeuGlyLysSerGlyPheSerLeuPheSerCysAlaTyrCysProGly--- 133
 QY 503 ATAATCTACCGCAAGAGCGCTTAGCTGCTCTCGGGTGTAGCTTGAGTTTGACAAAAAT 562
 Db 134 ---ThrThrGlyTyrGlyAlaIleGln---ThrLysGlyAsnThrThrLeuLysAspAsn 151
 QY 563 GTCAGTTTCTCTTACGCAAAACCTTTTCAACGGATATGCGGGTGTATCACGCCAAAA 622
 Db 152 SerSerLeuValPheHisLysAsnCysSerThrAlaGluGlyGlyAlaIleGlnCys--- 170
 QY 623 ACTCTTTTCATTAAACAGGAGTACAATGTCTGCTCTGTTTCTGAAAAATACCTCTCAAG 682
 Db 170 ----- 170
 QY 683 AAAGCGGAGCAATTCAGACTTCCGATGCCCTTACCATTTGGAACCAAGGGAAGTC 742
 Db 170 ----- 170
 QY 743 TCTTTTCTGACAATCTTCTTCGGATTCTGGAGTGCATTTTACAGAAAGCCTCGGTG 802
 Db 170 ----- 170
 QY 803 ACTATTTCTAATAATCTAAAGTTTCTTTTATTGACAATAAGGTACAGGAGGAGTCC 862
 Db 170 ----- 170
 QY 863 TCAACAACGGGGATATGTCTGAGGTGCTATCTGTCTTATAAACTAGTACAGTACT 922
 Db 171 -----LysGlySerSerAspAla 176
 QY 923 AAGTCAACCTCTACGAAATCAGATGTTACTCTTCAGCAACAATACATCGCAACAGCG 982
 Db 177 GluLeuLysIleGluAsnAsnGlnAsnLeuValPheSerGluAsnSerSerThrSerLys 196
 QY 983 GGAGGAGTACTATGTGAAAAAGCTCGAAGTTCGGAGGCTTCGGAGGACTTACCTTATCAGT 1042
 Db 197 GlyGlyAlaIleTyrAlaAspLysLeuThrIleValSerGlyGlyProThrLeuPheSer 216
 QY 1043 AGAAATAGTCTCAATGGAGGTACAGCTCTTAAGGTGGAGCCATGCTATCGAAGAT--- 1099
 Db 217 AsnAsnSerValSerAsnGlySerSerProLysGlyGlyAlaIleSerIleLysAspSer 236
 QY 1100 AGTGGGAATTTGATTTATCCCGCATAGTGGTGCATTTGTTTGGGAATACAGTC 1159
 Db 237 SerGlyGluCysSerLeuThrAlaAspLeuGlyAspIleThrPheAspGlyAsnLysIle 256
 QY 1160 ACTTCTACTACTCTCTGGGCGAAT-----AGAAAGTAGTATCGACTTAGGAAGCT 1210
 Db 257 IleLysThrSerGlyGlySerSerThrValThrArgAsnSerIleAspLeuGlyThr--- 275
 QY 1211 GCAAAGATGACAGCTTTGCTTCTGCTGTGTAGAGCCATCTACTTATGATCCATA 1270
 Db 276 GlyLysPheThrLysLeuArgAlaLysAspGlyPheGlyIlePhePheTyrAspProIle 295
 QY 1271 ACTACAGATCATCCACACAGTTCAGATGCTTTAAAGTTAATGAGACTCCGCGAGAT 1330
 Db 296 ThrGlyGlyGlySer-----AspGluLeuAsnIleAsnLys-----Lys 308
 QY 1331 TCTGCACTCAATATACAGGGAACATCTCTTACAGGAGAAAGTTATCAGACAGACAG 1390
 Db 309 GluThrValAspTyrThrGlyLysIleValPheSerGlyGluLysLeuSerAspGluGlu 328
 QY 1391 GCCGCAATTTCTAAATCTTACTTCTGAGCTACTACAGCTCTACTCTTTCAGAGGT 1450
 Db 329 LysAlaArgAlaGluAsnLeuAlaSerThrPheAsnGlnProIleThrLeuSerAlaGly 348
 QY 1451 ACTCTATCTTTAAACCATGAGGTGACTCTGCACTCAGGACTTCACTCAACAGGCGAGAT 1510

Db 349 SerLeuValLeuLysAspGlyValSerValThrAlaLysGlnValThrGlnGluAlaGly 368
QY 1511 TCTCGTCTCGAATGACGTAGCACTACTCTAGAACCT-----GCTGATACT 1558
Db 369 SerThrValValMetAspLeuGlyThrThrLeuGlnThrProSerSerGlyGlyGluThr 388
QY 1559 AGCACCATAACAATTTGGTCATTAACATCAGTCTTATA-----GACGGTGCAGAG 1609
Db 389 IleThrLeuThrAsnLeuAspIleAsnIleAlaSerLeuGlyGlyGlyThrSer 408
QY 1610 AAGCAGAAATAGAAACCAAGCTACGTCAAAAATCTGACTTACTTGGCAACCATCACT 1669
Db 409 ProAlaLysLeuAlaThrAsnThrAlaSerGlnAlaIleThrIleAsn---AlaValAsn 427
QY 1670 TTATTGGACCGGACGGCTTTTATGAAATCATAGTTTAAAGAAATCTCAGTCTCTAC 1729
Db 428 LeuValAspAlaAspGlyAsnAlaTyrGluAspProIleLeuAlaThrSerLysProPhe 447
QY 1730 GACATCTTAGAGCTCAAAAGCTTCTGGAACCTGTAAACAGCACCGAGTACTCCAGATCCT 1789
Db 448 ThrAlaIle-----ValAlaThrThrAsnAlaSerThrValThrGlnProThrAsp 464
QY 1790 ATAATGGGTGAG-----AAATTCATTACGGCTATCAGGGAACCTTGGGGCCCAATT 1840
Db 465 AsnLeuThrAsnTyrValProProThrHisTyrGlyThrGlnGlyAsnTrp---ThrVal 483
QY 1841 GTTTGGGGGACAGGGCTTCTAGC---ACTGCAACCTTCAACTGGACTAAACCTGGCTAT 1897
Db 484 ThrTrpAspThrGluThrAlaThrLysThrAlaThrLeuThrTrpGluGlnThrGlyTyr 503
QY 1898 ATTCTTAATCCGAGGTATCGGCTCTTATGCTCCTTAATAGCTTATGGAATGCTTTATA 1957
Db 504 SerProAsnProGluArgGlnGlyProLeuValProAsnThrLeuTrpGlyAlaPheSer 523
QY 1958 GATATTAGCTCTCCATTATCTTATGAGACTCAAAACGAAGGTTTCAGGGAGACCGT 2017
Db 524 AspLeuArgAlaIleGlnAsnLeuMetAspIleSerValAsnGlyAlaAspTyrHisArg 543
QY 2018 GCTTTTGGTGTGGTATATCTAATCTTCCATTAAGGATAGTACAAAACACGACGC 2077
Db 544 GlyPheTrpValSerGlyLeuAlaAsnPheLeuHisLysSerGlySerAspThrLysArg 563
QY 2078 GGGTTCCGCCATTGAGTGGCGGTATGCTATAGAGAAACCTACATCTTGTTCAGAT 2137
Db 564 LysPheArgHisAsnSerAlaGlyTyrAlaLeuGlyValTyrAlaLysThrProSerAsp 583
QY 2138 AAGATTCTTAGTCTGCATTTTGTGCTCTTTGGAAGAGATAGAGACTCTTTGTAGCT 2197
Db 584 AspIlePheSerAlaAlaPheCysGlnLeuPheGlyLysAspLysAspTyrLeuValSer 603
QY 2198 AAGATCAAGGTACAGTCTACGGAGGAACTCTCTATTACCAGCACACGAAACCTATATC 2257
Db 604 LysAsnAsnAlaAsnIleTyrAlaGlySerLeuTyrTyrGlnHis-----Ile 619
QY 2258 TCTCTCTCCGCAACTACGGCTTGTCTGTTTCTTATGTTCTCAGAGATTCCTGTT 2317
Db 620 SerTyrTrpSerAlaTrpGlnAsnLeuLeuGlnAsnThrIleGlyAlaGluAlaProLeu 639
QY 2318 CTCCTTTTCAGGAACTTAGCTACACCCATACGATACGATCTCAAAAC----- 2368
Db 640 ValLeuAsnAlaGlnLeuThrTyrCysHisAlaSerAsnAspMetLysThrAsnMetThr 659
QY 2369 -----AAGTATACACATATCTCTGTTTAAAGGAAGCTGGGGGAATGAT 2413
Db 660 ThrThrTyrAlaProArgLysThrThrTyrAlaGluIleLysGlyAspTrpGlyAsnAsp 679
QY 2414 AGTTTCGCTTTAGAATTCGGTGGAGAGCTCCGATTTGCTTAGATGAAGTCTCTATT 2473
Db 680 CysPheGlyValGluLeuGlyAlaThrValProIleGlnThrGluSerSerLeuLeuPhe 699
QY 2474 GAGCAGTACATCCCTTCATGAATTCAGTTTGTCTATGTCACATCAGGAAGTTTAA 2533
Db 700 AspMetTyrSerProPheLysPheGlnLeuValHisThrHisGlnAspAspPheLys 719

QY 2534 GAACAGGAGACA---GAAGCTCGTGAATTTGGAACTAGCGCTTGTGAATCTTGCTT 2590
Db 720 GluAsnAsnSerAspGlnGlyArgTyrPheGluSerSerAsnLeuThrAsnLeuSerLeu 739
QY 2591 CTTATCGGATCCGATTTGATAGGAATCAGACTGCCAAGATGCCAAGTACCAATCTAAC 2650
Db 740 ProIleGlyIleLysPheGluArgPheAlaAsnAsnAspThrAlaSerTyrHisValThr 759
QY 2651 CTTGGTTATPACTGTGATCTTGTTCGTAGTAACCCGACTGTACGACACACTGCGAAT 2710
Db 760 AlaAlaTyrSerProAspIleValArgSerAsnProAspCysThrThrSerLeuLeuVal 779
QY 2711 ACGCGTGATCT-----TGGAAACCTTCGGTACGAATTTGCAAGACAAGCTTTAGTC 2764
Db 780 SerProAspSerAlaValTrpValThrLysAlaAsnAsnLeuAlaArgSerAlaPheMet 799
QY 2765 CTTGCGTCAAGGAACCATTTTGTCTTAACTCAAAATTTGAAGCCTTTAGCCAATTTCT 2824
Db 800 LeuGlnAlaGlyAsnTyrLeuSerLeuSerHisAsnIleGluIlePheSerGlnPheGly 819
QY 2825 TTTGAATTCGGTGGTCTCTCTCGCAATTTACAATGTAGACTTAGGACGAAAATACCAATTC 2884
Db 820 PheGluLeuArgGlySerSerArgThrTyrAsnValAspLeuGlySerLysIleGlnPhe 839

RESULT 10

US-09-886-468-19
; Sequence 19, Application US/09886468
; Patent No. US20020037293A1
; GENERAL INFORMATION:
; APPLICANT: Aventis Pasteur Limited
; TITLE OF INVENTION: Chlamydia antigens and corresponding DNA fragments and uses therefor
; FILE REFERENCE: 77813-5
; CURRENT APPLICATION NUMBER: US/09/886,468
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/113,280
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113,281
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113,282
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113,283
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113,284
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113,285
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113,385
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/114,050
; PRIOR FILING DATE: 1998-12-28
; PRIOR APPLICATION NUMBER: 60/114,056
; PRIOR FILING DATE: 1998-12-28
; PRIOR APPLICATION NUMBER: 60/114,057
; PRIOR FILING DATE: 1998-12-28
; PRIOR APPLICATION NUMBER: 60/114,058
; PRIOR FILING DATE: 1998-12-28
; PRIOR APPLICATION NUMBER: 60/114,059
; PRIOR FILING DATE: 1998-12-28
; PRIOR APPLICATION NUMBER: 60/114,061
; PRIOR FILING DATE: 1998-12-28
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 19
; LENGTH: 922
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-886-468-19
Alignment Scores:
Pred. No.: 1.67e-109 Length: 922
Score: 1437.50 Matches: 350
Percent Similarity: 53.21% Conservative: 156

Best Local Similarity: 36.80%		Mismatches: 382	
Query Match: 26.58%		Indels: 63	
DB: 9		Gaps: 26	
US-09-428-122-1 (1-3000) x US-09-886-468-19 (1-922)			
QY	113	TTCCCC---AAGTTTGTATTTCTACATTTGCTATTTCCCTTTGTCTATGATGCTACC	169
Db	8	PheProLeuValPheSerPheThrLeuLeuSerValPheAspThrSerLeuSerAlaThr	27
QY	170	GAGACATTTTGGATTCAGTCGAGTTTCGATGGGAATAAAATGGTAATTTTCAGTT	229
Db	28	ThrIleSerLeuThrProGluAspSerPheHisGlyAspSerGlnAsn-----AlaGlu	45
QY	230	CGTGAGAGTCAGGAAGTCTCGAACTACCTACTCTATTTAAAGGAAATGCTCACTAGAA	289
Db	46	ArgSerTyrAsnValGlnAlaGlyAspValTyrSerLeuThrGlyAspValSerIleSer	65
QY	290	AATATTCCTGGAACAGGCACAGCAATCACAAAAGCTGTTTAAACAACACTAAGGGCGAT	349
Db	66	AsnVal-----AspAsnSerAlaLeuAsnLysAlaCysPheAsnValThrSerGlySer	83
QY	350	TTGACTTCACAGTAACGGAACTCTCTATTGTTTCCAAACGCTGGATCGAGGACTGTA	409
Db	84	ValThrPheAlaGlyAsnHisHisGlyLeuTyrPheAsnAsnIleSerSerGlyThrThr	103
QY	410	GCAGGGCTCTGTTTAAACAGCAGCGTGGPAGATAAATCTACCACG---TTTATAGGGTTT	466
Db	104	LysGluGlyAlaValLeuCysCysGlnAspProGlnAlaThrAlaArgPheSerGlyPhe	123
QY	467	TCCTTCGCTATCTTTTATTCGCTCTCTCGAAGTTCGATAACTACCGCAAGGAGCCGTT	526
Db	124	SerThrLeuSerPheIleGlnSerProGlyAspIleLysGluGlnGly-----	139
QY	527	AGCTGC-----TCTACGGGTAGCTTCAGTTTCACAAAATGCTCAGTTTCTCTTCAGC	580
Db	140	--CysLeuTyrSerLysAsnAlaLeuMetLeuLeuAsnAsnTyrValValArgPheGlu	158
QY	581	AAAACTTTTCAACGGNAATGCGGTGCTATCACCGCAAAACTCTTTCATTAAACAGGG	640
Db	159	GlnAsnGlnSerLysThrLysGlyGlyAlaIleSerGlyAlaAsnValThrIleValGly	178
QY	641	ACTACATGTCAGCTCTGTTTCTGAAATACCTCTCAAGAAAGCGGAGCCATTACAG	700
Db	179	AsnTyrAspSerValSerPheTyrGlnAsnAlaAlaThr---PheGlyGlyAlaIleHis	197
QY	701	ACTTCCGATGCCCTTACCATTTACGAAACCAAGGGGAAGTCTCTTTTCTGCAATACT	760
Db	198	SerSerGlyProLeuGlnIleAlaValAsnGlnAlaGluIleArgPheAlaGlnAsnThr	217
QY	761	TCCTCCGAT---TCTGGAGCTGCAATTTTACAGAGCCCTCGGTGACTATTCTTAATAAT	817
Db	218	AlaLysAsnGlySerGlyGlyAlaLeuTyrSerAspGlyAspIleAspIleAspGlnAsn	237
QY	818	GCTAAAGTTTCTTTTATTCACAAATAGGTCACAGGAGCGAGCTCCTCAACACGGGGAT	877
Db	238	AlaTyrValLeuPheArgGluAsnGlu-----AlaLeuThrThrAlaIle	252
QY	878	ATGTCAGGAGGTGCTATCTGTGCTTATAAACTAGTACAGATACTAAG-----	925
Db	253	GlyLysGlyGlyAlaValCysCysLeuProThrSerGlySerThrProValProIle	272
QY	926	GTCACCTCTACTGGAATTCAGATGTTACTCTTCAGCAACAATATCATCGAACACGGGGA	985
Db	273	ValThrPheSerAspAsnLysGlnLeuValPheGluArgAsnHisSerIleMetGlyGly	292
QY	986	GGAGCTATCTATGTGAAAGCTCGAACTGGCTTCGGAGGACTTACCTTATTTCAGTAGA	1045
Db	293	GlyAlaIleTyrAlaArgLysLeuSerIleSerSerGlyGlyProThrLeuPheIleAsn	312
QY	1046	AATAGTGTCAATGAGGTACAGTCTCTAAAGGTGGAGCCATAGCTATCGAAGATAGTGG	1105
Db	313	AsnIleSerTyrAlaAsnSerGlnAsnLeuGlyGlyAlaIleAlaIleAspThrGlyGly	332

QY	974	ACAACACGGGAGGAGCTATCTATGTGTAAGAAAGCTCGAACTGCTTCCGGAGAGCACTTACC	1033
Db	366	AlaGlyLysGlyGlyAlaIleTyAlaLysLysLeuSerValAlaAsnCysGlyProVal	385
QY	1034	CTATTTCAGTAGAATAAGTGTCAATGGAGGTACAGCTCTCTAAAGGTGGAGCCATAGCTATC	1093
Db	386	GlnPheLeuArgAsnIleAlaAsn-----AspGlyGlyAlaIleTyLeu	400
QY	1094	GAAGATAGTGGGAAATGAGTTTATCCGCGGATAGTGGTGACATGCTCTTTTAGGGAAT	1153
Db	401	GlyGluSerGlyGluLeuSerLeuSerAlaAspTyGlyAspIleIlePheAspGlyAsn	420
QY	1154	-----ACAGTCACTTCTACTCTCCCTGGGACGAAT-----AGAACT	1189
Db	421	LeuLysArgThrAlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGln	440
QY	1190	AGTATCACTTAGGAACGAGTGCAGAAAGATGACAGCTTTGCGTTCCTGCTGCTGAGGCC	1249
Db	441	AlaIleSerMetGlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGln	460
QY	1250	ATCTACTTCTATGATCCATAAATCACTCAGGATCATCCACAACA-----GTTACA	1297
Db	461	IleLeuPheAsnAspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSer	480
QY	1298	GATGTCTTAAAGATTAAAGACATCCCGGAGATTCGCATACATAATATACAGGAACATC	1357
Db	481	LysLeuLeuLysIleAsnAspGlyGluGly-----TyrThrGlyAspIle	495
QY	1358	ATCTTCACAGGAAAGATTATCAGACAGAGCGCCAGATTCCTAAAATCTTACTCTCG	1417
Db	496	ValPheIle-----AsnGlySerSer	502
QY	1418	AAGCTACTACAGCCTGTAACTCTTTCAGGAGTACTTAPCTTTTAAACATCGAGTCACT	1477
Db	503	ThrLeuTyGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLys	522
QY	1478	CTGCAGACTCAGGATTCATCAACACAGCAGATTCCTGCTCGAATGGAGCTAGGACT	1537
Db	523	LeuSerValAsnSerLeuSerGlnThrGlyGlySer---LeuTyThrMetGluAlaGlySer	541
QY	1538	ACTCTAGAA-----CCTGCTGATACATAGC-----	1561
Db	542	ThrLeuAspPheValThrProGlnProGlnGlnProProAlaAlaAsnGlnLeuIle	561
QY	1562	ACCATAACAATTTGGTCATTAAACATCAGTCTTATA-----	1597
Db	562	ThrLeuSerAsnLeuHisLeuSerLeuSerSerLeuLeuLeuAlaAsnAlaValThrAsn	581
QY	1598	-----GACGGTGCAAAGAGGCAAAATATAGAAACCAAGCTTACG	1636
Db	582	ProProThrAsnProProAlaGlnAspSerHisProAlaValIleGlySer---ThrThr	600
QY	1637	TCAAAANAATCTGACTTTATCTGGAACCATCATCTTATTTGGACCGCGGCACGTTTAT	1696
Db	601	AlaGlySerValThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTy	620
QY	1697	GAATAATCATACT---TTAAGAAATCCTCAGTCCCTACGACATCTTAGAGCTCAAAGCTTCT	1753
Db	621	AspArgTyAspTyrLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeu---	639
QY	1754	GGAAGTGTAAACAGCAGCGAGTCACTCCAGATCCTATATATGGTGAGAATTC---CAT	1810
Db	640	GlyThrLysProProAlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLys	659
QY	1811	TACGGCTATCAGGAACTTGGGGCCCAATCTTTTGGGGGACAGGGGCTTCTACAGCTGCA	1870
Db	660	TyrGlyTyGlnGlySerTrp---LysLeuAlaTyPaspProAsnThrAlaAsnAsnGly	678
QY	1871	ACCTTC-----AAGTGCATAAACTGGCTATATCTCTATCCGAGCGTATC	1918
Db	679	ProTyThrLeuLysAlaThrTrpThrLysThrGlyTyAsnProGlyProGluArgVal	698

Qy	1919	GGCTCTTTAGTCCTTAATAGCTTATATGAATGCAATTTATAGATATTAGCTCTCTCCCAATTAT	1970
Db	699	AlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSer	718
Qy	1979	CTTATGAGACTGCAACAAAGGGTTCGACGGGAGACCGTCTCTTTTGGTGTGCTGCATTAA	2038
Db	719	AlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyVal	738
Qy	2039	TCTAACTCTTTCATTAAGTAGTACAAAACACGACGCGGGTTTCGCCATTTGAGTGGC	2098
Db	739	SerAsnPhePheTyrHisAspArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGly	758
Qy	2099	GGTTATGCTACGAGGAACCTACATCTTGTTCAGATAAAGATCTTATAGTGTGCGATTT	2158
Db	759	GlyTyrSerLeuGlyAlaAsnSerTyrPheGlySer--SerMetPheGlyLeuAlaPhe	777
Qy	2159	TGTCAGCTCTTTGGAAGAGTAGAGACTACTTGTGTAGCTTAAGAATCAAGGTACAGTCTAC	2218
Db	778	ThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsnHisAlaCys	797
Qy	2219	GGAGGAACCTCTATTACCAGCACACGAAACCTATATCTCTCTCTGCAAACTACGG	2278
Db	798	IleGlySerValTyrLeuSerThrGlnGlnAlaLeu-----	809
Qy	2279	CCTTGTGTGCTCTTATATGTTCTCTACAGAGATCTCTGTCTCTTTTCAGGA-----	2329
Db	810	---CysGly---SerTyr-----LeuPheGlyAspAlaPheIle	820
Qy	2330	AACCTTAGCTACACCCATACGATAACGATCTGAAACCAAGATTAACAATATCCTACT	2389
Db	821	ArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGlu	840
Qy	2390	GTTAAAGGAAGCTGGGGAGTAGATGTTTCGCTTTAGAAATTCGGTGAAGAGCTCCGATT	2449
Db	841	SerAspValArgTrpAspAsnAsnCysLeuAlaGlyIleGlyAlaGlyLeuProIle	860
Qy	2450	TGCTTAGATGAAGTGTCTATT---GAGCAGTACATGCGCTTCATGAATTCGACGTTT	2506
Db	861	ValIleThrProSerLysLeuTyrLeuAsnGluLeuArgProPheValGlnAlaGluPhe	880
Qy	2507	GTCATGACATCAGGAAGGTTTAAAGACAGGGAACAGAGCTCGTGAATTTGGAAGT	2566
Db	881	SerTyrAlaAspHisGlySerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSer	900
Qy	2567	AGCCGCTCTGTGAATCTTGCTTACCTATCGGATCGATTTCATGAAGGAATCAGACTGC	2626
Db	901	GlyHisLeuLeuAsnLeuSerValProValGlyValLysPheAspArgCysSerSerThr	920
Qy	2627	CAAGATGCAAGTACATCTAATCTCTGTGGTTTACTGTGTGATCTGTTCGTAGTAACCC	2686
Db	921	HisProAsnLysTyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSer	940
Qy	2687	GACTGTACGACAAACACTCGCAATTAGCGGTGATCTTGGAAAAACCTTCGGTAGCAATTG	2746
Db	941	GlyThrGluThrThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeu	960
Qy	2747	GCAAGACAGCTTTAGTCTCTTCGTGACGGGAACATTTTGTCTTAACTCAAATTTTGAA	2806
Db	961	AlaArgHisGlyValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGlu	980
Qy	2807	GCCTTTAGCCCAATTTCTTTTGAATTCGGTGGGTCTCTCGCAATTAACAATGAGACTTA	2866
Db	981	ValTyrGlyHisGlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAla	1000
Qy	2867	GGAGCAAAATACCAATTC	2884
Db	1001	GlySerLysValArgPhe	1006

RESULT 12
US-09-841-132-176
; Sequence 176, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:


```

; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 176
; LENGTH: 982
; TYPE: PRT
; ORGANISM: Chlamydia
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(982)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-841-132-176

Alignment Scores:
Pred. No.: 9,89e-84 Length: 982
Score: 1124.50 Matches: 303
Percent Similarity: 45.92% Conservative: 159
Best Local Similarity: 30.12% Mismatches: 373
Query Match: 20.79% Indels: 171
DB: 9 Gaps: 30

US-09-428-122-1 (1-3000) x US-09-841-132-176 (1-982)
QY 242 GAAGATGCTGGAACTACCTACCTATTAAAGCGAAATGTCACTCTAGAAAATATTCTCGGA 301
Db 23 AppProSerGlyThrThrValPheSerAlaGlyGluLeuThrLeuLysLeuAspAsn 42
QY 302 ACAGGCACGCAATCAAAAAAGCTGTTTAAACACATAAGCGGCAATTGACTTTTACA 361
Db 43 SerIleAlaLeuProLeuSerCysPheGlyAsnLeuGlySerPheThrValLeu 62
QY 362 GGTAAAGGAACTCTATTGTTCCAAACGGTGGATGACGAGGACTGTACAGGGCTGCT 421
Db 63 GlyArgGlyHisSerLeuThrPheGluAsnIleArgThrSerThr---AsnGlyAlaAla 81
QY 422 GTTAAACAGCAGCGTGTAGATAAATCTACACAGCTTTATAGGGTTTCTTCGCTACTCTTT 481
Db 82 LeuSerAsnSerAlaAlaAspGlyLeuPheThrIleGluGlyPheLysGluLeuSerPhe 101
QY 482 -----ATTGCTCTCTCGAAGTTCGATTAACACCGGCAAGAGCCGTT 526
Db 102 SerAsnCysAsnSerLeuLeuAlaValLeuProAlaAlaThrThrAsnLysGlySerGln 121
QY 527 AGCTGCTCTACGGGTAGC-----TTGAGT 550
Db 122 ThrProThrThrThrSerThrProSerAsnGlyThrIleTyrSerLysThrAspLeuLeu 141
QY 551 TTGACAAAATGTCAGTTTGCTCTTCAGCAAAACTTTTCAACGGATAATGCGCGTGT 610
Db 142 LeuLeuAsnAsnGluLysPheSerPheTyrSerAsnLeuValSerGlyAspGlyGlyAla 161
QY 611 ATCAGCGCAAAACTCTTTCAATACAGGACTACAAGTCAAGTCAAGTCTGTTTCTGAAAT 670
Db 162 IleAspAlaLysSerLeuThrValGlnGlyIleSerLysLeuLysCysValPheGlnGluAsn 181
QY 671 ACTCTCTCAAGAAAGCGGCGCCATTTCAGACTTCGATTCGCTCCCTTACCATTACTGGAAC 730
Db 182 ThrAlaGlnAlaAspGlyGlyAlaCysGlnValValThrSerPheSerAlaMetAlaAsn 201
QY 731 CAAGGGGAAGTCTCTTT----- 748
Db 202 GluAlaProIleAlaPheValAlaAsnValAlaGlyValArgGlyGlyGlyIleAlaAla 221
QY 749 -----TGTGAATACTCTCTCGATTCT----- 772
Db 222 ValGlnAspGlyGlnGlnGlyValSerSerThrSerThrSerThrGluAspProValValSer 241

```

```

QY 773 -----GGAGCTCAATT 784
Db 242 PheSerArgAsnThrAlaValGluPheAspGlyAsnValAlaArgValGlyGlyIle 261
QY 785 TTTACAGAAAGCTCGTGACTATTCTTAATATGCTAAAGTTTCTTTATTACCAAT--- 841
Db 262 TyrSerTyrGlyAsnValAlaPheLeuAsnGlyLysThrLeuPheLeuAsnVal 281
QY 842 -----AAGGTCACAGGAGCGAGTCTCTCAACAAG 871
Db 282 AlaSerProValTyrIleAlaAlaLysGlnProThrSerGlyGlnAlaSerAsnThrSer 301
QY 872 GGGGATATG---TCAGGAGGTCTATCTGTGCTTATAA-----ACTAGT 913
Db 302 AsnAsnTyrGlyAspGlyGlyAlaIlePheCysLysAsnGlyAlaGlnAlaGlySerAsn 321
QY 914 ACAGATACTAAGGTCACCTCACTCGAAATCAGATGTTACTCTTCAGCAACAATACATCG 973
Db 322 AsnSerGlySerValSerPheAspGlyGluGlyValValPhePheSerSerAsnValAla 341
QY 974 ACAACAGCGGAGGAGCTATCTATGTGAAAGCTCGAACTGGCTTCGAGGAGCTTACC 1033
Db 342 AlaGlyLysGlyGlyAlaIleTyrAlaLysLysLeuSerValAlaAsnCysGlyProVal 361
QY 1034 CTATTTCAGTAGAATAGTGTCAATGAGGTACAGCTCTCTAAAGGTGGAGCCATGCTATC 1093
Db 362 GlnPheLeuArgAsnIleAlaAsn-----AspGlyGlyAlaIleTyrLeu 376
QY 1094 GAAGATAGTGGGAAATTCAGTTTATCCCGCATAGTGTGATGATGCTTTTACGGAAAT 1153
Db 377 GlyGluSerGlyGluLeuSerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsn 396
QY 1154 -----ACAGTCACCTTCTACTCTCTGGAGCAAT-----AGAAGT 1189
Db 397 LeuLysArgThrAlaLysGluAsnAlaAlaAspValAsnGlyValThrValSerSerGln 416
QY 1190 AGTATCGACTAGGACGAGTGCAAGATGACAGCTTTGCGTTCTGCTGCTGGTAGAGCC 1249
Db 417 AlaIleSerMetGlySerGlyGlyLysIleThrThrLeuArgAlaLysAlaGlyHisGln 436
QY 1250 ATCTACTTCTATGATCCCACTAATACAGATCATCCACAACA-----GTTACA 1297
Db 437 IleLeuPheAsnAspProIleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSer 456
QY 1298 GATGCTTTAAAGTTAATGAGACTCCGCGAGATTCGCACATCAATATACAGGGAACATC 1357
Db 457 LysLeuLeuLysIleAsnAspGlyGly-----TyrThrGlyAspIle 471
QY 1358 ATCTTCACAGGAGAAAGTTATACAGACAGCGCGCAGATTCTTAAAAATCTTACTTCG 1417
Db 472 ValPheAla-----AsnGlySerSer 478
QY 1418 AAGTACTACAGCTGTAACTCTTTACAGAGTACTCTATCTTTTAAAACATGGAGTGACT 1477
Db 479 ThrLeuTyrGlnAsnValThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLys 498
QY 1478 CTGACAGCTAGGCAATTCACCAACAGCGAGATTCGTCTCGAAATAGCAGTAGGAAT 1537
Db 499 LeuSerValAsnSerLeuSerGlnThrGlyGlySer---LeuTyrMetGluAlaGlySer 517
QY 1538 ACTCTAGAA-----CCTGCTGATACATAGC----- 1561
Db 518 ThrLeuAspPheValThrProGlnProGlnProProAlaAlaAsnGlnLeuIle 537
QY 1562 ACCATAACAATTTGGTCAATTAACATCAGTTCTATA----- 1597
Db 538 ThrLeuSerAsnLeuHisLeuSerLeuSerLeuLeuAlaAsnAsnAlaValThrAsn 557
QY 1598 -----GACGGTGCAGAGGCAAAAATAGAAACCAAGCTACG 1636
Db 558 ProProThrAsnProProAlaGlnAspSerHisProAlaValIleGlySer---ThrThr 576

```



```
QY 1637 TCAAAATCTGACTTATCTGGAACCATCACATTATTGACCGCGGCAGCTTTTAT 1696
   ::::::::::::::::::::|||
Db 577 AlaGlySerValThrIleSerGlyProIlePhePheGluAspLeuAspAspThrAlaTyr 596
QY 1697 GAAATCATAGT---TTAAGAAATCCCTCAGTCTTACGACATCTTAGAGCTCAAGCTTCT 1753
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 597 AspArgTyrAspTrpLeuGlySerAsnGlnLysIleAsnValLeuLysLeuGlnLeu--- 615
QY 1754 GGAACCTGTAACAAGCACCGCAGTGACTCCAGATCCTATAATGGTGAGAAATTC---CAT 1810
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 616 GlyThrLysProProAlaAsnAlaProSerAspLeuThrLeuGlyAsnGluMetProLys 635
QY 1811 TAGGGCTATCAGGAACCTGGGGCCCAATGTGTGGGGACAGGGGCTTCTACAGCTGCA 1870
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 636 TyrGlyTyrGlnGlySerTrp---LysLeuAlaTrpAspProAsnThrAlaAsnAsnGly 654
QY 1871 ACCTTC-----AATCGACTAAACCTGGCTATATCTCTAATCCGACGCGTATC 1918
   ::::::::::::::::::::|||
Db 655 ProTyrThrLeuLysAlaThrTrpThrLysThrGlyTyrAsnProGlyProGluArgVal 674
QY 1919 GGCTCTTTAGTCCCTAATAGCTTATGAATGATCATTTATAGATATAGCTCTCTCCATAT 1978
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 675 AlaSerLeuValProAsnSerLeuTrpGlySerIleLeuAspIleArgSerAlaHisSer 694
QY 1979 CTTATGAGACTGCAACAGGAGGTTGCAGGAGACGCTGCTTTTGGTGTGCTGATTA 2038
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 695 AlaIleGlnAlaSerValAspGlyArgSerTyrCysArgGlyLeuTrpValSerGlyVal 714
QY 2039 TCTAATCTTCTCCATAAGGATAGTACAAAACACGACGCGGTTTCGCCATTTGAGTGGC 2098
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 715 SerAsnPhePheTyrHisAspArgAspAlaLeuGlnGlyTyrArgTyrIleSerGly 734
QY 2099 GGTATGTCATAGAGGAACCTACATCTTGTTCAGATAAGATCTTAGTGTGCAATTT 2158
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 735 GlyTyrSerLeuGlyAlaAsnSerTyrPheGlySer---SerMetPheGlyLeuAlaPhe 753
QY 2159 TGTGAGCTCTTGGAGAGATAGAGACTCTTTCTAGCTAAGATCAAGGTACAGTCTAC 2218
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 754 ThrGluValPheGlyArgSerLysAspTyrValValCysArgSerAsnHisHisAlaCys 773
QY 2219 GGAGGAATCTCTATTACCAGCACAAACGAAACCTATATCTCTCTCTGCAAACTACGG 2278
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 774 IleGlySerValTyrLeuSerThrGlnGlnAlaLeu----- 785
QY 2279 CTTTGTGCTGTCTTATGTTCTTACAGAGATTCCTGTCTCTTTTCAGGA----- 2329
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 786 ---CysGly---SerTyr-----LeuPheGlyAspAlaPheIle 796
QY 2330 AACCTTAGCTACCCCATACGATAACGATCTGAAACCAAGTATACACATATCTTACT 2389
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 797 ArgAlaSerTyrGlyPheGlyAsnGlnHisMetLysThrSerTyrThrPheAlaGluGlu 816
QY 2390 GTTAAAGGAAGCTGGGGAATGATGTTTCGCTTTTAGAATTCGGTGGGAAGCTCCGATT 2449
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 817 SerAspValArgTrpAspAsnAsnCysLeuAlaGlyIleGlyAlaGlyLeuProIle 836
QY 2450 TGCTTAGTAAGATGCTCTATT---GAGCAGTACATGCCCTTCATGAAATTCAGTTT 2506
   ::::::::::::::::::::|||
Db 837 ValIleThrProSerLysLeuTyrLeuAsnGlnLeuArgProPheValGlnAlaGluPhe 856
QY 2507 GTCTATGCACATCAGGAAGGTTTAAAGAACAGGGAACAGAGCTCGTGAATTTGGAAGT 2566
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 857 SerTyrAlaAspHisGlySerPheThrGluGluGlyAspGlnAlaArgAlaPheLysSer 876
QY 2567 AGCGCTCTTGTGAATCTTTCCTATCCGATCCGATCCGATTTGATGAAATTCAGACTGC 2626
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 877 GlyHisLeuLeuAsnLeuSerValProValGlyValLysPheAspArgCysSerSerThr 896
QY 2627 CAAGATGCAACGTACATCTTCTGTTTACTGTGATCTTGTGTTAGTAAACCC 2686
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 897 HisProAsnLysTyrSerPheMetAlaAlaTyrIleCysAspAlaTyrArgThrIleSer 916
QY 2687 GACTGTACACAACTCGGAATAGCGGTGATCTTGGAAAACTTCGGTACCAATTTG 2746
```

```
Db 917 GlyThrGluThrThrLeuLeuSerHisGlnGluThrTrpThrThrAspAlaPheHisLeu 936
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 2747 GCAAGACAAGCTTTAGTCTCTCGTCAGGGAACCATTTTTCCTTTAACTCAAAATTTTGAA 2806
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 937 AlaArgHisGlyValValValArgGlySerMetTyrAlaSerLeuThrSerAsnIleGlu 956
QY 2807 GCCTTTAGCAATTTTCTTTGAATTCGTTGGTGCATCTCGCAATTTACGAATGTAGACTTA 2866
   ::::::::::::::::::::|||
Db 957 ValTyrGlyHisGlyArgTyrGluTyrArgAspAlaSerArgGlyTyrGlyLeuSerAla 976
QY 2867 GGAGCAAAATACCAATTC 2884
Db 977 GlySerLysVal***Phe 982

RESULT 13
US-09-841-132-175
; Sequence 175, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 175
; LENGTH: 880
; TYPE: PRT
; ORGANISM: Chlamydia
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(880)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-841-132-175

Alignment Scores:
Pred. No.: 2,19e-62 Length: 880
Score: 865.00 Matches: 262
Percent Similarity: 44.51% Conservative: 123
Best Local Similarity: 30.29% Mismatches: 366
Query Match: 15.99% Indels: 114
DB: 9 Gaps: 24

US-09-428-122-1 (1-3000) x US-09-841-132-175 (1-880)
QY 452 ACGTTTATAGGTTTCTTCGCTATCTTTATTGCGTCTCTCGAAGTTCGATAACTACC 511
   ::::::::::|||
Db 76 SerTyrCysTrpPheValSerLysLeuHisIleThrAspProLysGluAlaLeuPheLys 95
QY 512 GCGAAAGGAGCGCTTAGC-----TGCTCTACG 538
   |||||:::|||||
Db 96 GluLysGlyAspLeuSerIleGlnAsnPheArgPheLeuSerPheThrAspCysSerSer 115
QY 539 -----GGTAGCTTGAGTTTGACAAAATAAT 562
   ::::::::::|||
Db 116 LysGluSerSerProSerIleIleHisGlnLysAsnGlyGlnLeuSerLeuArgAsnAsn 135
QY 563 GTCAGTTTGTCTTCAGCAAAACTTTTCAACGATAATGCGGTGCTATCACCGCAAAA 622
   ::::::::::|||
Db 136 GlySerMetSerPheCysArgAsnHisAlaGluGlySerGlyGlyAlaIleSerAlaAsp 155
QY 623 ACTCTTTTCATTACAGGACTACATGTCAGCTCTGTTTCTGAAATACCTCTCAAG 682
   |||||:::|||||
Db 156 AlaPheSerLeuGlnHisAsnTyrLeuPheThrAlaPheGluGluAsnSerSerLysGly 175
QY 683 AAAGGCGGAGCCATTTCAGACTTCGATGCCCTTACCATTACTGAAACCAAGGGAGTC 742
   |||||:::|||||
Db 176 AsnGlyGlyAlaIleGln---AlaGlnThrPheSerLeuSerArgAsnValSerProIle 194
```

QY 743 TCTTTTCTGACAAATCTCTTCGGATTCTCGAGCTGCAATTTTACAGAGCCTCGGTG 802
Db 195 SerPheAlaArgAsnArgAlaAspLeuAsnGlyGlyAlaIleCysSerAsnLeuIle 214
QY 803 ACTATTCTTAATAATGCTAAAGTTTCTCTTTATTGACAATAAGTTCACAGAGCGAGCTCC 862
Db 215 CysSerGlyAsnValAsnProLeuPhePheThrGlyAsnSerAlaThr 230
QY 863 TCAACAACGGGGGATATGTCAGAGGTGCTATCTGTGCT-----TATRAAACTAGT 913
Db 231 -----AsnGlyGlyAlaIleCysCysIleSerAspLeuAsnThrSer 244
QY 914 ACAGATACTAAGTCACTACCCTCACTGGAATCAGATGTTACTCTTCAGCAACAATACATCG 973
Db 245 GluLysGlySerLeuSerLeuAlaCysAsnGlnGluThrLeuPheAlaSerAsnSerAla 264
QY 974 ACAACACGGGGAGGAGCTATCTATGTGAAAAGCTCGAACTGGCTTCCGGAGGACTTACC 1033
Db 265 LysGluLysGlyAlaIleTyrAlaLysHisMetValLeuArgTyrAsnGlyProVal 284
QY 1034 CTATTCACTAGATAATAGTGTCAATGAGGTACAGCTCTCTAAAGTGGAGCCATAGCTATC 1093
Db 285 SerPheIleAsnAsnSer-----AlaLysIleGlyAlaIleAlaIle 299
QY 1094 GAAGATAGTGGGAATGAGTTTATCCGCCGATAGTGTGACATGCTCTTTTAGGGAAT 1153
Db 300 GlnSerGlyLysSerLeuSerIleLeuAlaGlyGluGlySerValLeuPheGlnAsnAsn 319
QY 1154 ACAGTCACTTCTACTACTCTCGGGACGAATAGAAAGTAGTATCGACTTAGGAACGAGTGCA 1213
Db 320 SerGlnArgThrSerAspGlnGlyLeuValArgAsnAlaIleTyrLeu**LysAspAla 339
QY 1214 AGATGACAGCTTTGGTCTGCTGCTGTGTGAGGCCATCTACTCTTATGATCCCAACT 1273
Db 340 IleLeuSerSerLeuGluAlaArgAsnGly--AspIleLeuPhePheAspProIleVal 358
QY 1274 ACAGGATCATCCACACAGTTACAGATGTCTTAAAGTTAATGAGACTCGGCAGATTCT 1333
Db 359 GlnGluSerSerSer-----LysGluSerProLeuProSer 370
QY 1334 GCACATAAATATACA----- 1348
Db 371 SerLeuGlnAlaSerValThrSerProThrProAlaThrAlaSerProLeuValIleGln 390
QY 1349 -----GGGAACATCATCTTCACAGGAGAAAAGTTATCAGAGACAGAGCCGCA 1396
Db 391 ThrSerAlaAsnArgSerValIlePheSerSerGluArgLeuSerGluGluLysThr 410
QY 1397 GATTCTAAAACTCTTACTTCTGAGCTACTACAGCCTGTAACTCTTTTCAGAGGACTCTTA 1456
Db 411 ProAsp---AsnLeuThrSerGlnLeuGlnProIleGluLeuLysSerGlyArgLeu 429
QY 1457 TCTTTAAACATGGAGTCTCTGACAGCTCAGGCATTCACCTCAACAGGCAGATTCTCGT 1516
Db 430 ValLeuLysAspArgAlaValLeuSerAlaProSerLeuSerGlnAspProGlnAlaLeu 449
QY 1517 CTCGAAATCGAGTACAGACTCTCTAGAACCTCTGCTAGTACACCACTAAACAATTTG 1576
Db 450 LeuIleMetGluAlaGlyThrSerLeuLysThrSerSerAspLeuLysLeuAlaThrLeu 469
QY 1577 GTCATTAACTCAGTCTATAGACGTGCAGAGCGGCAAGAGGCAAAATAGAAACCAAGCTACG 1636
Db 470 SerIleProLeuHisSerLeuAspThrGluLysSerValThrIleHis----- 485
QY 1637 TCAAAAATCTGACTTTTATCTGGAACCATCACTTTATTGGACCGGCGGACGTTTAT 1696
Db 486 AlaProAsnLeuSerIleGlnLysIlePheLeuSerAsnSerGlyAspGluAsnPheTyr 505
QY 1697 GAAATCATAGTTTAAAGAACTCTCAGTCTACGACATCTTAGACTCAAGCTTCTGGA 1756
Db 506 GluAsnValGluLeuLeuSerLysGluGlnAsnAsnIle---ProLeuLeuThrLeuPro 524

QY 1757 ACTGTAAACAGCACCGCAGTGACTCCAGATCTCTATAATGGGTGAGAAAATCCATTACGCC 1816
Db 525 LysGluGlnSerHisLeuHisLeuProAspGlyAsnLeuSerSer-----HisPheGly 542
QY 1817 TATCAGGGAACTTGGGGCCCAATTGTTTGG-----GGGACAGGGCTTCTACGACT 1867
Db 543 TyrGlnGlyAspTrp---ThrPheSerTrpLysAspSerAspGluGlyHisSerLeuIle 561
QY 1868 GCAACCTTCAACTGGACTAAAACTCGCTATATTCCTTAATCCCGAGCGTATCGCTCTTTA 1927
Db 562 Ala-----AsnTrpThrProLysAsnTyrValProHisProGluArgGlnSerThrLeu 579
QY 1928 GTCCCTAATAGCTTATGGAATGCATTTATAGATATTAGCTCTCTCCATTATCTTATGAG 1987
Db 580 ValAlaAsnThrLeuTrpAsnThrTyrSerAspMetGlnAlaValGlnSerMetIleAsn 599
QY 1988 ACTGCAAAACGAAGGTTGCAGGAGACCGTGTCTTT-----TGTGTGCTGATATCT 2041
Db 600 ThrThrAlaHisGly-----GlyAlaTyrLeuPheGlyThrTrpGlySerAlaValSer 617
QY 2042 AACTTCTTC-----CATAAAGGATAGTACAAAACACACGCGGGTTTCGCCATTGAGT 2095
Db 618 AsnLeuPheTyrValHisAspSerSerGlyLysProIleAspAsnTrpHisArgSer 637
QY 2096 GCGGCTTATGCTAGGAGGAAACCTACATCTTGTTCAGATAAGATTCTTGTGCTGCA 2155
Db 638 LeuGlyTyrLeuPheGlyIleSerThrHisSerLeuAspHisSerPheCysLeuAla 657
QY 2156 TTTTGTGCTCTTTTGGAGAGATAGAGACTACTTTTGTAGCTAAAGATCAAGGTACAGTC 2215
Db 658 AlaGlyGlnLeuLeuGlyLysSerSerAspSerPheIleThrSerThrGluThrThrSer 677
QY 2216 TACGAGGAACTCTCTATTACGACACACAGAACTATATCTCTCTCTCTTGCACAACTA 2275
Db 678 TyrIleAlaThrValGlnAlaGlnLeuAlaThrSerLeuMetLysIle----- 693
QY 2276 CGGCTTGTCTGCTTATGTTCTCTACAGAGATCTGTTCTCTTTTCAGAAACCTT 2335
Db 694 -----SerAlaGlnAla 697
QY 2336 AGCTACCCCATACGATAACGATCTGAAAACCAAGTATACAACTATCTACTGTAA 2395
Db 698 CysTyrAsnGluSerIleHisGluLeuLysThrLysTyrArgSerPheSerLysGluGly 717
QY 2396 ---GGAAGTGGGGAATGATAGTTTCGCTTTTAGAATTCGGTGGAGAGCTCCGATTTGC 2452
Db 718 PheGlySerTrpHisSerValAlaValSerGlyGluValCysAlaSerIleProIleVal 737
QY 2453 TTAGATGAAGTCTCTATTGAGCAGTACATGCCCTTCATGAAATTGCAGTTTGTCTAT 2512
Db 738 SerAsnGlySerGlyLeuPheSerPheSerIlePheSerLysLeuGlnGlyPheSer 757
QY 2513 GCACATCAGAAAGTGTAAAGAACAGGACAGAGCTCGTGAATTTTGGAGTAGCCGT 2572
Db 758 GlyThrGlnAspGlyPheGluGluSerSerGlyGluIleArgSerPheSerAlaSerSer 777
QY 2573 CTTTGAATCTTGCCTTACCTATCGGATCCGATTGTAGTAAAGATCAGACTGCCAAGAT 2632
Db 778 PheArgAsnIleSerLeuProIleGlyIleThrPheGluLysLysSerGln---LysThr 796
QY 2633 GCACAGTACATCTAACTCTTGGT---TATCTGTGGATCTTGTGCTGATGTAACCCGAC 2689
Db 797 ArgThrTyrTyrPheLeuGlyAlaTyrIleGlnAspLeuLysArgAspValGluSer 816
QY 2690 TGTACGACAACTCCGAATTAGCGGTGATCTCTCGAAAACCTTCGGTAGCAATTTGGCA 2749
Db 817 GlyProValValLeuLeuLysAsnAlaValSerTrpAspAlaProMetAlaAsnLeuAsp 836
QY 2750 AGACAAGCTTTAGTCTCTTCGTGCGAGGAAACCACTTTTGTCTTAACTCAAAATTTGAAGC 2809
Db 837 SerArgAlaTyrMetPheArgLeuThrAsnGlnArgAlaLeuHis---ArgLeuGlnThr 855
QY 2810 TTTACCAATTTTCTTTTGAATTCGGTGGTCAATCTCGCAATTACAATGTAGACTTAGGA 2869

```

Db      856 LeuLeuAsnValSerCysValLeuArgGlyGlnSerHisSerThrLeuAspLeuGly 875
QY      2870 GCAAAATACCAATTC 2884
Db      876 ThrThrTyArgPhe 880

RESULT 14
US-09-841-132-189
; Sequence 189, Application US/09841132
; Patent No. US2002061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yaser A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841.132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 189
; LENGTH: 866
; TYPE: PRT
; ORGANISM: Chlamydia
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(866)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-841-132-189

Alignment Scores:
Pred. No.: 1,2e-61 Length: 866
Score: 856.00 Matches: 260
Percent Similarity: 44.28% Conservative: 123
Best Local Similarity: 30.06% Mismatches: 368
Query Match: 15.83% Indels: 114
DB: 9 Gaps: 24

US-09-428-122-1 (1-3000) x US-09-841-132-189 (1-866)
QY      452 ACGTTTATAGGTTTCTTCGCTATCTTTATTCGGCTCTCGAAGTCGATTAACATACC 511
Db      62 SerTyrcystrpPheValSerLysLeuHisIleThrAspProLysGluAlaLeuPheLys 81
QY      512 GGCAAGGAGCGGTAGC-----TGCTCTACG 538
Db      82 GluLysGlyAspLeuSerIleGlnAsnPheArgPheLeuSerPheThrAspCysSerSer 101
QY      539 -----GGTAGCTTGAGTTTGACAAAAAT 562
Db      102 LysGluSerSerProSerIleIleHisGlnLysAsnGlyGlnLeuSerLeuArgAsnAsn 121
QY      563 GTCAGTTGCTCTCAGCAAAACTTTCAACGATATATGCGGTCTATCCGCGCAAAA 622
Db      122 GlySerMetSerPheCysArgAsnHisAlaGluGlySerGlyAlaIleSerAlaAsp 141
QY      623 ACTCTTTCATTACAGGAGCTACAAATGTCAGCTCTGTTTCTGAAATACCTCTCAAG 682
Db      142 AlaPheSerLeuGlnHisAsnTyrrLeuPheThrAlaPheGluGluAsnSerSerLysGly 161
QY      683 AAAGCGGAGCCATTACAGATCCGATCCCTTACCATTACTGAAACCAAGGGGAATC 742
Db      162 AsnGlyGlyAlaIleGln--AlaGlnThrPheSerLeuSerArgAsnValSerProIle 180
QY      743 TCTTTTCTGCAATATCTTCTCGGATTCGAGCTGCAATTTTACAGACCTCGGTG 802
Db      181 SerPheAlaArgAsnArgAlaAspLeuAsnGlyGlyAlaIleCysCysSerAsnLeuIle 200
QY      803 ACTATTCTTAATAGCTAAAGTTCTCTTATTGACATAAGGTCACAGGAGCGAGCTCC 862
Db      201 CysSerGlyAsnValAsnProLeuPhePheThrGlyAsnSerAlaThr----- 216

```

```

QY      863 TCAACAACGGGGATATGTCACGAGGTGCTATCTGTGCT-----TATAAACTAGT 913
Db      217 -----AsnGlyGly***IleCysCysIleSerAspLeuAsnThrSer 230
QY      914 ACAGATACTAAGGTCACCTCAGTGAATACATGTTACTCTTCAGCAACAATATACG 973
Db      231 GluLysGlySerLeuSerLeuAlaCysAsnGln***ThrLeuPheAlaSerAsnSerAla 250
QY      974 ACAACAGCGGAGGAGCTATCTATGTGAAAAAGCTCGAACTGGCTTCGCGAGGACTTACC 1033
Db      251 LysGluLysGlyGlyAlaIleTyrrAlaLysHisMetValLeuArgTyrrAsnGlyProVal 270
QY      1034 CTATTTCAGTAGAAATAGTCAATGGAGGTACAGCTCTTAAGAGTGGAGCATAGTATC 1093
Db      271 SerPheIleAsnAsnSer-----AlaLysIleGlyGlyAlaIleAlaIle 285
QY      1094 GAAGATAGTGGGAATTGAGTTTATCCGCGATAGTGGTACATTTGTCATTTTAGGGAAT 1153
Db      286 GlnSerGlyGlySerLeuSerIleLeuAlaGlyGluGlySerValLeuPheGlnAsnAsn 305
QY      1154 ACAGTCACTTCTACTCTCTCGGACGAATAGAAAGTAGTAGTACGACTTAGAACGAGTCA 1213
Db      306 SerGlnArgThrSerAspGlnGlyLeuValArgAsnAlaIleTyrrLeuGluLysAspAla 325
QY      1214 AAGATGACAGCTTTGCGTCTGCTGTGTAGAGCCATCTACTTCTATGATCCATAACT 1273
Db      326 IleLeuSerSerLeuGluAlaArgAsnGly--AspIleLeuPhePheAspProIleVal 344
QY      1274 ACAGGATCATCCACACAGTACAGATGCTTAAAGTTAATGAGACTCGGCGAGATTCT 1333
Db      345 GlnGluSerSerSer-----LysGluSerProLeuProSer 356
QY      1334 GCACATAAATATACA----- 1348
Db      357 SerLeuGlnAlaSerValThrSerProThrProAlaThrAlaSerProLeuValIleGln 376
QY      1349 -----GGGAACATCATCTTCACAGAGAGAAAGTTATCAGAGACAGAGCGGCA 1396
Db      377 ThrSerAlaAsnArgSerValIlePheSerSerGluArgLeuSerGluGluLysThr 396
QY      1397 GATTCTAAAATCTTACTTCGAGCTACTACAGCTGTAACTCTTTTCAGAGAGTACTCTA 1456
Db      397 ProAsp---AsnLeuThrSerGlnLeuGlnGlnProIleGlnLeuLysSerGlyArgLeu 415
QY      1457 TCTTTTAAAACATCGAGTGACTCTCGACTCAGCACTTCACCAAGGAGGAGATTCTCGT 1516
Db      416 ValLeuLysAspArgAlaValLeuSer***ProSerLeuSerGlnAspProGlnAlaLeu 435
QY      1517 CTCGAATGACGTAGGAACACTCTTAGAACCTGTGATAGTACCACTAACAATTTG 1576
Db      436 LeuIleMetGluAlaGlyThrSerLeuLysThrSer***AspLeuLysLeu***Thr*** 455
QY      1577 GTCATTAAACATCAGTCTATAGACGGTGCAGAGGCAAGGCAAAATAGAACCAAGCTACG 1636
Db      456 SerIleProLeuHisSerLeuAspThrGluLysSerValThrIleHis----- 471
QY      1637 TCAAAAATCTGACTTTTATCTGGAACCATCTATTATGGACCCGCGGACGCTTTAT 1696
Db      472 AlaProAsnLeuSerIleGlnLysIlePheLeuSerAsnSerGlyAspGluAsnPheTyrr 491
QY      1697 GAAATCATAGTTTAGAAATCTCTAGTCTCTAGTCTCTAGTCTTAGAGCTTCAAGCTTCGGA 1756
Db      492 GluAsnValGluLeuSerLysGluGlnAsnAsnIle---ProLeuLeuThrLeuPro 510
QY      1757 ACTGTAACAGACCCGAGTCACTCCAGATCTCTATATGGTGGTGAATAATCCATTACGGC 1816
Db      511 LysGluGlnSerHisLeuHisLeuProAspGlyAsnLeuSerSer-----HisPheGly 528
QY      1817 TATCAGGGAACCTTGGGCGCCCAATTGTTTGG-----GGGACAGGCGCTTCTACGACT 1867
Db      529 TyrGlnGlyAspTrp---ThrPheSerTrpLysAspSerAspGluGlyHisSerLeuIle 547

```

1868	QY	GCACCTTCACTGCATAAACCTGGCTATATCTCTAATCCCGAGCGTATCGGCTCTTTA	1927
548	Db	Ala-----AsnThrPrProLysAsnTyrValProHisProGluArgGlnSerThrLeu	565
1928	QY	GTCCCTAATAGCTTATGAATGCATTTATACATATTAGCTCTCCATTATCTATGAG	1987
566	Db	ValAlaIasnThrLeuTPAsnThrTyrSerAspMetGlnAlaValGlnSerMetIleAsn	595
1998	QY	ACTGCAACGAAGGGTTGCAGGAGACCGTCTTTT-----TGGTGTCTGGATATCT	2041
586	Db	ThrThrAlaHisGly-----GlyAlaTyrLeuPheGlyThrTyrGlySerAlaValSer	603
2042	QY	AAC TTCCTC-----CATACGATAGTACAAAACACGACGCGGGTTCCGCCATTTGAGT	2095
604	Db	AsnLeuPheTyrValHisAspSerSerGlyLysProIleAspAsnThrHisLeuArgSer	623
2096	QY	GGCGGTTATGTCATAGGAGGAAACCTACATCTTTGTCAGATAGATCTTGTGCTGCA	2155
624	Db	LeuGlyTyrLeuPheGlyIleSerThrHisSerLeuAspAspHisSerPheCysLeuAla	643
2156	QY	TTTTGTGAGCTCTTTGGNAGAGATAGAGACTTTGTAGCTTAGAATCAAGGTACAGTC	2215
644	Db	AlaGlyGlnLeuLeuGlyLysSerSerAspSerPheIleThrSerThrGluThrThrSer	663
2216	QY	TACGGAGGAACCTCTCTATTACCAACCAACGAAACCTATATCTCTCTCTCTGCAAACTA	2275
664	Db	TyrIleAlaThrValGlnAlaGlnLeuAlaThrSerLeuMetLysIle-----	679
2276	QY	CGGCTTGTGCTGTCTTAGTTGTTCTACAGAGATTCCTGTCTCTTTTCAGGAAACCTT	2335
680	Db	-----SerAlaGlnAla	683
2336	QY	AGCTACACCCATACGGATAACGATCTGAAACCAAGTATACACATATCCTACTGTATAA	2395
684	Db	CysTyrAsnGluSerIleHisGluLeuLysThrLysTyrArgSerPheSerLysGluGly	703
2396	QY	---GGAAGCTGGGGAAATGATGTTTCGTTTAGAATTCGGTGGGAAGCTCCGATTTGC	2452
704	Db	PheGlySerThrHisSerValAlaValSerGlyGluValCysAlaSerIleProIleVal	723
2453	QY	TTAGTCAAAAGTCTCTATTGTGACGATCATGCCCCTTCATGAAATTCGAGTTTGCTAT	2512
724	Db	SerAsnGlySerGlyLeuPheSerSerPheSerIlePheSerLysLeuGlnGlyPheSer	743
2513	QY	GCACATCAGAAAGGTTTTAAAGAACACGGGAACAGAGCTCGTGAATTTGGAAGTACCGT	2572
744	Db	GlyThrGlnAspGlyPheGluGluSerSerGlyGluIleArgSerPheSerAlaSerSer	763
2573	QY	CTTGTGAATCTGCTTACCTATCTCGGATCCGATTTGATAAAGAACTCAGACTGCCAAGT	2632
764	Db	PheArgAsnIleSerLeuProIleGlyIleThrPheGluLysLysSerGln---LysThr	782
2633	QY	GCAACGTACAAATCTAACTCTTGT---TATACGTGATCTCTCTCGTAGTACACCCGAC	2689
783	Db	ArgThrTyrTyrPheLeuGlyAlaTyrIleGlnAspLeuLysArgAspValGluSer	802
2690	QY	TGTACGCAACACTGCGAATTAGCGGTGATCTCTGGAAAACTTCGGTACGAATTTGGCA	2749
803	Db	GlyProValValLeuLeuLysAsnAlaValSerTripAspAlaProMetAlaAsnLeuAsp	822
2750	QY	AGCAGAGCTTTAGTCTCTCGTCGAGGGAACCATTTTGTCTTAACTCAATTTTGAAGCC	2809
823	Db	SerArgAlaTyrMetPheArgLeuThrAsnGlnArgAlaLeuHis---ArgLeuGlnThr	841
2810	QY	TTTAGCCAATTTCTTTTGAATTCGCTGGGTGATCTCGCAATPACAAATGACATPAGGA	2869
842	Db	LeuLeuAsnValSerCysValLeuArgGlyGlnSerHisSerTyrSerLeuAspLeuGly	861
2870	QY	GCAAAATACCAATTC	2884
862	Db	ThrThrTyrArgPhe	866

RESULT 15

```

US-09-738-269-53
; Sequence 53, Application US/09738269
; Publication No. US20030185948A1
; GENERAL INFORMATION:
; APPLICANT: JOHNSTON, STEPHEN A.
; APPLICANT: STEMKE-HALE, KATHERINE
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: KALTENBOECK, BERNHARD
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF
; FILE REFERENCE: CHLAMYDIA PSITTACI
; FILE REFERENCE: UTSD:659
; CURRENT APPLICATION NUMBER: US/09/738,269
; CURRENT FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 679
; TYPE: PRT
; ORGANISM: Chlamydia psittaci
US-09-738-269-53

Alignment Scores:
Pred. NO.: 2.61e-61
Score: 851.50
Percent Similarity: 45.62%
Best Local Similarity: 30.88%
Query Match: 15.74%
DB: 12
Gaps: 23
Matches: 243
Conservative: 116
Mismatch: 283
Indels: 145

```

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 16, 2003, 10:21:34 ; Search time 48.5 Seconds

(without alignments)
11897.157 Million cell updates/sec

Title: US-09-428-122-1

Perfect score: 5409

Sequence: 1 cgcctctacacgtagaggt.....tgctttgctaaacactttc 3000

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 566616

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09428122/runat_16122003_102128_10789/app_query.fasta_1.3143
-DB=PIR_76 -QFMT=fastan -SUFFIX=trp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=oct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09428122@cgn_1_1_75 @runat_16122003_102128_10789 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_76.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3991	73.8	772	2	H86492
2	2058	38.0	928	2	G86546
3	2058	38.0	928	2	G81591
4	1982	36.6	928	2	E86546
5	1982	36.6	928	2	E72077
6	1965	36.3	936	2	C72078
7	1964	36.3	936	2	C86546
8	1964	36.3	936	2	B81591
9	1940	35.9	930	2	D86546
10	1940	35.9	930	2	A81591
11	1936	35.8	930	2	D72078
12	1868	34.5	949	2	F81591
13	1855	34.3	928	2	H86546
14	1855	34.3	928	2	D72077

15	1580.5	29.2	841	2	E72130
16	1444	26.7	1276	2	B86546
17	1444	26.7	1276	2	C81591
18	1442.5	26.7	922	2	B72131
19	1442.5	26.7	922	2	E86491
20	1441.5	26.7	922	2	F81539
21	1384.5	25.6	995	2	C81593
22	1383.5	25.6	1407	2	B72078
23	1377.5	25.5	973	2	B86547
24	1377.5	25.5	973	2	F72076
25	1253.5	23.2	712	2	E86492
26	1120.5	20.7	1013	2	G71460
27	1051	19.4	987	2	H81722
28	1017.5	18.8	445	2	E86493
29	917	17.0	186	2	G86492
30	916	16.9	359	2	C86493
31	911	16.8	867	2	F81721
32	863	16.0	878	2	B71460
33	821	15.2	494	2	D86493
34	792	14.6	427	2	A86493
35	685.5	12.7	1609	2	A86611
36	685.5	12.7	1609	2	G72013
37	672.5	12.4	978	2	G72076
38	670	12.4	947	2	G86557
39	667.5	12.3	978	2	B81593
40	667.5	12.3	978	2	C86547
41	666	12.3	947	2	D72067
42	666	12.3	1723	2	H86557
43	666	12.3	1723	2	E72067
44	666	12.3	1732	2	C81601
45	665.5	12.3	946	2	C86549

ALIGNMENTS

RESULT 1

H86492
Pmp_3 [imported] - Chlamydomophila pneumoniae (strain J138)
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
R:Accession: H86492
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Ouchi, K.; Shiba, T.;
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: H86492
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-772 <STO>
A:Cross-references: GB:BA000008; NID:G8978389; PIDN:BAA98226.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: pmp_3_2

Alignment Scores:
Pred. No.: 8.16e-245 Length: 772
Score: 3991.00 Matches: 771
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.87% Mismatches: 0
Query Match: 73.78% Indels: 0
DB: 2 Gaps: 0

US-09-428-122-1 (1-3000) x H86492 (1-772)

QY	569	TTGCTCTTCAGCAAAACCTTTTCAACGGTAATGCGGTGCTATCACCGCAAAACTCTT	628
DB	1	MetLeuPheSerLysAsnPheSerThrAspAsnGlyGlyAlaIleThrAlaLysThrIeu	20
QY	629	TCATTAAACAGGAGCTACAATCTGCTGCTGTTTCTGAAAATACCTCTCAAGAAAGGC	688
DB	21	SerLeuThrGlyThrThrMetSerAlaLeuPheSerGluAsnThrSerSerLysLysGly	40
QY	699	GGAGCCATTACACTTCGGATGCCCTTACCTACTTGAAACCAAGGGAGTCTCTTTT	748

Db 41 GlyAlaIleGlnThrSerAspAlaLeuThrIleThrGlyAsnGlnGlyLeuValSerPhe 60
QY 749 TCTGACAAATCTCTTCCGATTCGAGCTGCAAATTTTACAGAAAGCCTCGGTGACTATT 808
Db 61 SerAspAsnThrSerSerAspSerGlyAlaAlaIlePheThrGluAlaSerValThrIle 80
QY 809 TCTAATAATGCTAAAGTTCTCTTATTGACAAATAGGTCTCAGAGAGGAGCTCTCAACA 868
Db 81 SerAsnAsnAlaLysValSerPheIleAspAsnLysValThrGlyAlaSerSerThr 100
QY 869 ACGGGGATATGTCAGAGGTGCTATCTGCTTATATAAACTAGTACAGATACCTAAGGTC 928
Db 101 ThrGlyAspMetSerGlyGlyAlaIleCysAlaIleThrSerThrAspThrLysVal 120
QY 929 ACCCTCACTGGAATCAGATGTTACTCTTCAGCAACAATCATCGACAAACAGCGGGAGGA 988
Db 121 ThrLeuThrGlyAsnGlnMetLeuLeuPheSerAsnAsnThrSerThrAlaGlyGly 140
QY 989 GCTATCTATGTGAAAAGCTCGAACTCGCTTCGGAGGACTTACCTATTTCAGTAGAAAT 1048
Db 141 AlaIleThrValLysLysLeuGluLeuAlaSerGlyGlyLeuThrLeuPheSerArgAsn 160
QY 1049 AGTGTCAATGGAGGTACAGCTCTCTAAAGGTGGAGCCATAGCTATCGAAGATAGTGGGAA 1108
Db 161 SerValAsnGlyGlyThrAlaProLysGlyGlyAlaIleAlaIleGluAspSerGlyGlu 180
QY 1109 TTGAGTTTATCCCGCATGATGTTGACATGCTCTTTTGGGAATACAGTCATCTCTACT 1168
Db 181 LeuSerLeuSerAlaAspSerGlyAspIleValPheLeuGlyAsnThrValThrSerThr 200
QY 1169 ACTCTGGGACGAAATAGAGTATGATCGACTTAGGACGAGTCAAGATGACAGCTTTG 1228
Db 201 ThrProGlyThrAsnArgSerSerIleAspLeuGlyThrSerAlaLysMetThrAlaLeu 220
QY 1229 CGTCTGCTCTGGTAGAGCCATCTACTTCTATGATCCCACTACAGGATCATCCACA 1288
Db 221 ArgSerAlaAlaGlyArgAlaIleThrPheThrAspProIleThrThrGlySerSerThr 240
QY 1289 ACAGTTACAGATGCTTAAAGTTAATGACATCCGCGACATTCGACATCAATATACA 1348
Db 241 ThrValThrAspValLeuLysValAsnGluThrProAlaAspSerAlaLeuGlnThr 260
QY 1349 GGGAACTATCTTCAGGAGAAAGTATATCAGACAGACGCGCAGATTTCTAAAT 1408
Db 261 GlyAsnIleIlePheThrGlyGlyLysLeuSerGluThrGluAlaAlaAspSerLysAsn 280
QY 1409 CTTACTTCGAAGCTACTACAGCCTGTAACTCTTTCAGAGGTACTCTATCTTTAAACAT 1468
Db 281 LeuThrSerLysLeuLeuGlnProValThrLeuSerGlyGlyThrLeuSerLeuLysHis 300
QY 1469 GGAGTGACTCTGCAGACTCGGATTCACCTCAACAGCAGATTCGCTCGAAATGGAC 1528
Db 301 GlyValThrLeuGlnThrGlnAlaPheThrGlnGlnAlaAspSerArgLeuGluMetAsp 320
QY 1529 GTAGGAATCTCTAGAACCTGTGTACTAGACACATAAACAATTTGGTCATTAAACATC 1588
Db 321 ValGlyThrThrLeuGluProAlaAspThrSerThrIleAsnAsnLeuValIleAsnIle 340
QY 1589 AGTTCTATAGACGGTGCAAGAGGCAAAATAGAAACCAAGCTACGTCACAAAATCTG 1648
Db 341 SerSerIleAspGlyAlaLysLysAlaLysIleGluThrLysAlaThrSerLysAsnLeu 360
QY 1649 ACTTTATCTGGAACCATCTTTATGACCCGACCGGACGCTTTTATGAAATCATAGT 1708
Db 361 ThrLeuSerGlyThrIleThrLeuLeuAspProThrGlyThrPheThrGluAsnHisSer 380
QY 1709 TTAAGAAATCTCAGTCTCAGCATCTTAGAGCTCAAGCTCTCAAGCTGCAAGCTGAAACAGC 1768
Db 381 LeuArgAsnProGlnSerThrAspIleLeuGluLeuLysAlaSerGlyThrValThrSer 400
QY 1769 ACCGAGTACTCCAGATCTCTAATGGGTGAGAAATTCATTTACGGCTATCAGGGAAT 1828

Db 401 ThrAlaValThrProAspProIleMetGlyGlyLysPheHisThrGlyThrGlnGlyThr 420
QY 1829 TGGGGCCCAATTTGTTGGGGACAGGGGCTTTACAGCTGCACCTTCAACTCGACTAAA 1888
Db 421 TripGlyProIleValThrGlyThrGlyAlaSerThrThrAlaThrPheAsnThrThrLys 440
QY 1889 ACTGGCTATATCTCTAATCCCGAGCGCTATCGGCTCTTTAGTCCCTAATAGCTTATGGAAT 1948
Db 441 ThrGlyThrIleProAsnProGluArgIleGlySerLeuValProAsnSerLeuThrAsn 460
QY 1949 GCATTTATAGATATAGCTCTCTCCATTTATCGAGAGTCGACAAACGAGGTTGCGAG 2008
Db 461 AlaPheIleAspIleSerSerLeuHisThrLeuMetGluThrAlaAsnGluGlyLeuGln 480
QY 2009 GGAGACGGTGTCTTTTGGTGTCTGATTTACTTCTTCCATAAGGATAGTACAAA 2068
Db 481 GlyAspArgAlaPheThrCysAlaGlyLeuSerAsnPheHisLysAspSerThrLys 500
QY 2069 ACACGACGGGTTTCCGCAATTTGAGTGGGCTTATGATCATAGGAGAAACCTACTACT 2128
Db 501 ThrArgArgGlyPheArgHisLeuSerGlyGlyThrValIleGlyAsnLeuHisThr 520
QY 2129 TGTTCAGATAAGATTTCTAGTCTGCTATTTGCTAGCTCTTTGGAAGAGATAGACTAC 2188
Db 521 CysSerAspLysIleLeuSerAlaAlaPheCysGlnLeuPheGlyArgAspArgSerThr 540
QY 2189 TTTGTAGCTAAGAAATCAAGTACAGTCTACGGAGGAACCTCTTATTCACGACACACGAA 2248
Db 541 PheValAlaLysAsnGlnGlyThrValThrGlyGlyThrLeuThrThrGlnHisAsnGlu 560
QY 2249 ACCTATATCTCTCTCTGCAAACTACGGCTTGTCTGCTTATGTTCTATGTTCTACAGAG 2308
Db 561 ThrThrIleSerLeuProCysLysLeuArgProCysSerLeuSerThrValProThrGlu 580
QY 2309 ATTCTGTGTTCTCTTTTCAGAAACCTTAGCTACACCATACGGATAACCATCTGAAAAC 2368
Db 581 IleProValLeuPheSerGlyAsnLeuSerThrHisThrAspAsnAspLeuLysThr 600
QY 2369 AGTATACAACTATCTCTACTGTTAAAGAGCTGGGGAATGATAGTTTCGCTTTAGAA 2428
Db 601 LysThrThrThrThrProThrValLysGlySerThrGlyAsnAspSerPheAlaLeuGlu 620
QY 2429 TTCGGTGAAGAGCTCCGATTTCTTAGATGAAGTCTCTATTTTTCAGCAGTACATGCC 2488
Db 621 PheGlyGlyArgAlaProIleCysLeuAspGluSerAlaLeuPheGluGlnThrMetPro 640
QY 2489 TTCATGAATTCGAGTTTGTCTATGCACATCAGGAGGTTTAAAGAACAGGACACAGAA 2548
Db 641 PheMetLysLeuGlnPheValThrAlaHisGlnGlyGlyPheLysGluGlnGlyThrGlu 560
QY 2549 GCTCGTGAATTTGGAAGTAGCGCTCTTGTGAATCTTTCCTTACCTATCGGATTCGATTT 2608
Db 661 AlaArgGluPheGlySerSerArgLeuValAsnLeuAlaLeuProIleGlyIleArgPhe 680
QY 2609 GATAAGGAATCAGACTGCCAAGATCAAGTACATCTAATCTCTTGTGTTATCTGAT 2668
Db 681 AspLysGluSerAspCysGlnAspAlaThrThrAsnLeuThrLeuGlyThrValAsp 700
QY 2669 CTTGTTCTGATTAACCCGACTGTACGACCACTCGCAATTTAGCGGTGATTTCTGGAAA 2728
Db 701 LeuValArgSerAsnProAspCysThrThrThrLeuArgIleSerGlyAspSerThrLys 720
QY 2729 ACCTTCGGTACCAATTTGGCAAGACAGCTTTAGTCTCTTCGTCGAGGGAACCATTTTGC 2788
Db 721 ThrPheGlyThrAsnLeuAlaArgGlnAlaLeuValLeuArgAlaGlyAsnHisPheCys 740
QY 2789 TTTACTCAATTTTGAAGCTTTAGCCAAATTTCTTTGATTTGCTGGTGGTCACTCCG 2848
Db 741 PheAsnSerAsnPheGluAlaPheSerGlnPheSerPheGluLeuArgGlySerSerArg 760
QY 2849 AATTACAATGTAGACTTAGGACAAAATACCAATTC 2884
Db 761 AsnThrAsnValAspLeuGlyAlaLysThrGlnPhe 772


```
QY 1799 GAGAAATTCCTTACGGCTATCAGGGAACCTTGGGCCCCAATGTGTTGG----- 1846
Db 564 ThrProThrHisTyrGlyTyrGlnGlyThrTrpGly---MetThrTrpValAspThr 582
QY 1847 ---GGGACAGGGGCTTCTACGACTGCACCTTCAACTGGACTAAACCTGGCTATATTCCT 1903
Db 583 AlaSerThrProLysThrLysThrAlaThrLeuAlaThrAsnThrGlyTyrLeuPro 602
QY 1904 AATCCCGAGCGTATPCGGCTCTTTAGTCCCTTAATAGCTTATGGAATGCATTTATAGATATT 1963
Db 603 AsnProGluArgGlnGlyProLeuValProAsnSerLeuTrpGlySerPheSerAspIle 622
QY 1964 AGCTCTCTCATATCTTATGGAGACTGCACAAACGAGGTTGCAGGAGACCGTCTTTT 2023
Db 623 GlnAlaIleGlnGlyValIleGluArgSerAlaLeuThrLeuCysSerAspArgGlyPhe 642
QY 2024 TGGTGTCTGTGATATCTTCTTCCATAAGGATAGTACAAAACACACGCGGTTT 2083
Db 643 TrpAlaAlaGlyValAlaAsnPheLeuAspLysAspLysGlyGluLysAArgLysTyr 662
QY 2084 CGCCATTGAGTGGCGTTATGTCATAGGAGGAAACCTACATACTTGTTCAGATAGATT 2143
Db 663 ArgHisLysSerGlyGlyTyrAlaIleGlyAlaAlaGlnThrCysSerGluAsnLeu 682
QY 2144 CTTAGTCTGCATTTTCTCAGCTCTTTGGAGAGATAGACTACTTTGTAGCTAGAAT 2203
Db 683 IleSerPheAlaPheCysGlnLeuPheGlySerAspLysAspPheLeuValAlaLysAsn 702
QY 2204 CAAGGTACAGCTACGGAGGAACCTCTTATACCAGCACACGAAACCTATATCTCTT 2263
Db 703 HisThrAspThrTyrAlaGlyAlaPheTyrIleGlnHis-----IleThrGlu 718
QY 2264 CTTTGCAAACTACGGCTCTGTTCTGTTCTTATGTTCTCT-----ACAGAGATTCC 2314
Db 719 CysSerGlyPheIleGlyCysLeuLeuAspLysLeuProGlySerTrpSerHisLysPro 738
QY 2315 GTTCTCTTTTCAGGAACCTTAGCTACCCATACGATACCGATACCATCTGAAAAACAGTAT 2374
Db 739 LeuValLeuGluGlnLeuAlaTyrSerHisValSerAsnAspLeuLysThrLysTyr 758
QY 2375 ACAACATATCTACTGTAAAGGAGCTGGGGGAATGATAGTTTTCGTTTGAATTCGGT 2434
Db 759 ThrAlaTyrProGluValLysGlySerTrpGlyAsnAlaPheAsnMetMetLeuGly 778
QY 2435 GGAAGAGCTCCGATTGCTTAGTAGAAGAGCTCTATTGAGCAGTACATGCCCTTCATG 2494
Db 779 AlaSerSerHisSerTyrProGluTyrLeuHisCysPheAspThrTyrAlaProTyrIle 798
QY 2495 AAATTGCAAGTTTCTCTATGCACATCAGGAAGTTTAAAGAACAGGACAGAGCTCGT 2554
Db 799 LysLeuAsnLeuThrTyrIleArgGlnAspSerPheSerGluLysGlyThrGluGlyArg 818
QY 2555 GAATTTGGAAGTAGCCGCTCTTGTGAATCTTGCCTTACCTATCCGGATCCGATTGATAAG 2614
Db 819 SerPheAspAspSerAsnLeuPheAsnLeuSerLeuProIleGlyValLysPheGluLys 838
QY 2615 GAATCAGACTGCCAAGATGCAAGTCAATCTAACTCTTGGTTATCTGTGGATCTGTT 2674
Db 839 PheSerAspCysAsnAspPheSerTyrAspLeuThrLeuSerTyrValProAspLeuIle 858
QY 2675 CGTAGTAACCCGACTGTAGCAACACACTCCGAATTAGCGGTGATCTTGGAAAACTTC 2734
Db 859 ArgAsnAspProLysCysThrThrAlaLeuValIleSerGlyAlaSerTrpGluThrTyr 878
QY 2735 GSTACGAATTTGGCAAGCAAGCTTTAGTCTCTTCGTGCAGGGAACCATTTTGTCTTAAC 2794
Db 879 AlaAsnAsnLeuAlaArgGlnAlaLeuGlnValArgAlaGlySerHisTyrAlaPheSer 898
QY 2795 TCMAATTTTGAAGCCCTTTAGCCAAATTTCTTTTGAATTCGTTGGGTCACTCGCAATTAC 2854
Db 899 ProMetPheGluValLeuGlyGlnPheValPheGluValArgGlySerSerArgIleTyr 918
```

```
QY 2855 AATGTAGACTTAGGAGCAAAATACCAATTC 2884
Db 919 AsnValAspLeuGlyGlyLysPheGlnPhe 928
RESULT 3
G81591
polymorphic membrane protein G family CP0303 [imported] - Chlamydomphila pneumoniae (s
C;Species: Chlamydomphila pneumoniae, Chlamydia pneumoniae
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 15-Jun-2001
Accession: G81591
R;Read: T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: G81591
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-928 <REA>
A;Cross-references: GB:AE002192; GB:AE002161; NID:g7189226; PIDN:AAF38160.1; PID:g718
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: CP0303
C;Superfamily: Chlamydomphila pneumoniae polymorphic outer membrane protein G
Alignment Scores:
Pred. No.: 1,64e-122 Length: 928
Score: 2058.00 Matches: 444
Percent Similarity: 61.68% Conservative: 142
Best Local Similarity: 46.74% Mismatches: 320
Query Match: 38.05% Indels: 44
DB: 20 Gaps: 20
US-09-428-122-1 (1-3000) x G81591 (1-928)
QY 101 ATGAGTCTCTCTTCCCAAGTTTGTATTT---TCTACATTTGCTATTTTC----- 148
Db 1 MetLysSerGlnPheSerTrpLeuValLeuSerSerThrLeuAlaCysPheThrSerCys 20
QY 149 CTTTGTCTATGATGTCTACCGAGACAGTTTGGATTCAAGTCCGAGTTTCGATCGGAAT 208
Db 21 SerThrValPheAlaAlaThrAlaGluAsnIleGlyProSerAspSerPheAspGlySer 40
QY 209 AAAAAAT---GGTAATTTTTCAGTTCTGTAGAGTCAAGAGATGCTGGAACTACCTACCTA 265
Db 41 ThrAsnThrGlyThrTyrThrProLysAsnThr-----ThrThrGlyIleAspTyrThr 58
QY 266 TTTAAGGGAAATGTCACCTCTAGAAAATATTCTCGAACAGGCACAGCAATCAAAAAGC 325
Db 59 LeuThrGlyAspIleThrLeuGlnAsnLeu---GlyAspSerAlaAlaLeuThrLysGly 77
QY 326 TGTTTAAACACACTAAGGGCGATTGACTTTTACAGGTAACGGGAACCTCTATTGTTTC 385
Db 78 CysPheSerAspThrThrGluSerLeuSerPheAlaGlyLysGlyTyrSerLeuSerPhe 97
QY 386 CAAACGGTGGATGACGAGGACTGTAGACGGGCTGCTGTGTAAACAGCAGCGTGGTAGATAA 445
Db 98 LeuAsnIleLysSer---SerAlaGluGlyAlaAlaLeu---SerValThrThrAspLys 115
QY 446 TCTACCAGTTTATAGGTTTCTTTCGTTATCTTTTATTCGCTCTCTCGGAAGTTCGATA 505
Db 116 AsnLeuSerLeuThrGlyPheSerSerLeuThrPheLeuAlaAlaProSerSerValIle 135
QY 506 ACTAC-----GGCAAGGAGCCCTTAGCTCTACGGGTAGCTTCAGTTTGACAAA 559
Db 136 ThrThrProSerGlyLysGlyAlaValLysCysGly---GlyAspLeuThrPheAspAsn 154
QY 560 AATGTCAGTTGCTCTTCAGCAAAAACCTTTCAACGGATAATGGCGGTCTACTACCGCA 619
Db 155 AsnGlyThrIleLeuPheLysGlnAspTyrCysGluGluAsnGlyGlyAlaIleSerThr 174
QY 620 AAAACTCTTTTCATTAAAGGAGCACTCAATGTCAGCTCTGTTTCTGAAAAATACCTCTCA 679
```


Db 819 SerPheAspSerGlyAspLeuIleAsnValSerValProIleGlyThrPheGluArg 838
 QY 2615 GATCAGACTGCCAGATGCAAGTCAATCTACTTCTGGTTATCTGATCTGTT 2674
 Db 839 PheSerArgAsnGluArgAlaSerTyrGluAlaThrValIleTyrValAlaAspValTyr 858
 QY 2675 CGTAGTAACCCGACTGTACGACACACATCGCAATTAGCGGTGATCTTGGAAACCTTC 2734
 Db 859 ArgLysAsnProAspCysThrThrAlaLeuLeuLeuAsnThrSerTyrPheThrThr 878
 QY 2735 GGTACGAATTGGCAAGACAGCTTGTAGTCTTCTGCGACGGGAACCATTTTGTCTTAAC 2794
 Db 879 GlyThrAsnLeuSerArgGlnAlaGlyIleGlyArgAlaGlyIlePheTyrAlaPheSer 898
 QY 2795 TCRAATTTTCAAGCCTTTAGCCAAATTTCTTTTGAATTCGGTGGTCACTCGCAATTAC 2854
 Db 899 ProAsnLeuGluValThrSerAsnLeuSerMetGluIleArgGlySerSerArgSerTyr 918
 QY 2855 AATGTAGACTTAGGAGCAAAATACCAATTC 2884
 Db 919 AsnAlaAspLeuGlyGlyLysPheGlnPhe 928
 RESULT 6
 C72078
 polypeptide outer membrane protein g family - Chlamydothila pneumoniae (strain CWL029)
 C:Species: Chlamydothila pneumoniae, Chlamydia pneumoniae
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 15-Jun-2001
 C:Accession: C72078
 R:Kalmann, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
 Nature Genet. 21, 385-389, 1999
 A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
 A:Reference number: A72000; MUID:99206606; PMID:10192388
 A:Accession: C72078
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-936 <ARN>
 A:Cross-references: GB:AE001627; GB:AE001363; NID:g4376721; PIDN:AAID18589.1; PID:g437672
 A:Experimental source: strain CWL029
 C:Genetics:
 A:Gene: pmp 7
 C:Superfamily: Chlamydothila pneumoniae polypeptide outer membrane protein G
 Alignment Scores:
 Pred. No.: 1.26e-116 Length: 936
 Score: 1965.00 Matches: 411
 Percent Similarity: 62.30% Conservative: 179
 Best Local Similarity: 43.40% Mismatches: 327
 Query Match: 36.33% Indels: 30
 DB: 2 Gaps: 16
 US-09-428-122-1 (1-3000) x C72078 (1-936)
 QY 101 ATGAGTCTCTTCCCAAGTTGTATTTTACATTTGCTATTTTC---CCTTTGTCT 157
 Db 1 MetLysSerSerValSerTyrPhePheSerSerLeuSer 20
 QY 158 ATGATGTCTACCGACAGATTTTGGATTCAAGT---CGGAGTTTCGATGGGAATAAAT 214
 Db 21 IleValAlaAlaGluValThrLeuAspSerSerAsnAsnSerTyrAspGlySerAsnGly 40
 QY 215 GGTAATTTTTCAGTTTCGTCGAGTCAGGAA---GATGTCGAACTACCTACCTATTAAAG 271
 Db 41 ThrThrPheThrValPheSerThrThrAspAlaAlaGlyThrThrTyrSerLeuLeu 60
 QY 272 GGAAATGTCACTCTAGAAAATATTCTCGAACAGGACAGCAATCAACAAAAGCTGTTTT 331
 Db 61 SerAspValSerPheGlnAsnAlaGlyAlaLeuGlyIleProLeuAlaSerGlyCysPhe 80
 QY 332 AACAACTAAGCGGATTTGACCTTTACAGTTAACGGGAACCTCTATTGTTTCCAAAG 391
 Db 81 LeuGluAlaGlyGlyAspLeuThrPheGlnGlyAsnGlnHisAlaLeuLysPheAlaPhe 100

QY 392 GTGATGACGAGGACTGTAGCAGGGCTGCTGTTTAAACAGCAGCGTGGTAGATAAATCTACC 451
 Db 101 IleAsnAlaGlySerSerAlaGlyThrValAlaSerThrSerAlaAlaAspLysAsnLeu 120
 QY 452 ACGTTTATAGGTTTCTTTCGCTATCTTTTATTCGCTCTCTCGAAGTTCGATAACT--- 508
 Db 121 LeuPheAsnAspPheSerArgLeuSerIleAsnLysCysProSerLeuLeuLeuSerPro 140
 QY 509 ACCGCGAAAGGAGCGGTAGCTGTCTACGGGTAGCTTGTAGTTGTACAAAAATGTCCAGT 568
 Db 141 ThrGlyGlnCysAlaLeuLys---SerValGlyAsnLeuSerLeuThrGlyAsnSerGln 159
 QY 569 TTGCTCTTACGAAAAAATCTTTTCAACGGATTAAGCGCGTCTATACCGCAAAATCTCTT 628
 Db 160 IleIlePheThrGlnAsnPheSerSerAspAsnGlyGlyValIleAsnThrLysAsnPhe 179
 QY 629 TCATTACAGGAGTACAAATGTCTGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 682
 Db 180 LeuLeuSerGlyThrSerGlnPheAlaSerPheSerArgAsnGlnAlaPheThrGlyLys 199
 QY 683 AAAGCGGAGCCATTACAGACTTCCGATGCCCTTACCAATTACGAAACCAAGGGAAGTC 742
 Db 200 GlnGlyGlyValValTyrAlaThrGlyThrIleThrIleGluAsnSerProGlyIleVal 219
 QY 743 TCTTTTCTGACATACTTCTTCCGATCTCGAGCTGCAATTTTACAGAGCTCGGTG 802
 Db 220 SerPheSerGlnAsnLeuAlaLysGlySerGlyAlaLeuTyrSerThrAspAsnCys 239
 QY 803 ACTATTTCTAATAATGCTAAAGTTTCTTATTGCAATAAGGTACAGAGGAGGAGCTCC 862
 Db 240 SerIleThrAspAsnPheGlnValIlePheAspGlyAsnSerAlaTyrGluAlaAlaGln 259
 QY 863 TCAACACGGGGATATGTACAGAGGTCTATCTGCTCTTATAAATAGTACAGATACT 922
 Db 260 Ala-----GlnGlyGlyAlaIleCysCys-----ThrThrThrAspLys 272
 QY 923 AAGTTCACCTCCTGGAATACAGATGTTCTCTTACAGCAACATATACATACAGCAACAGCG 982
 Db 273 ThrValThrLeuThrGlyAsnLysAsnLeuSerPheThrAsnAsnThrAlaLeuThrTyr 292
 QY 983 GGAGGAGCTATCTATGTGAAAGCTCGAATCTTCCGAGGAGCTTACCTATTTCAGT 1042
 Db 293 GlyGlyAlaIleSerGlyLeuLysValSerIleSerAlaGlyGlyProThrLeuPheGln 312
 QY 1043 AGAATAGTCTCAATGAGGTACAGCTCTCTAA-----GGTGGAGCATAGCTATCGAA 1096
 Db 313 SerAsn---IleSerGlySerSerAlaGlyGlnGlyGlyGlyGlyAlaIleAsnIleAla 331
 QY 1097 GATAGTGGGGAATGAGTTTATCCGCGATAGTGTGACATTGTCTTTTAGGGAATACA 1156
 Db 332 SerAlaGlyGluLeuAlaLeuSerAlaThrSerGlyAspIleThrPheAsnAsnAsnGln 351
 QY 1157 GTCATCTTACTACTCTCTGGGACGAATAGATAGTATCGATAGGACGAGTGGCAAG 1216
 Db 352 ValThrAsnGlySerThrSerThr---ArgAsnAlaIleAsnIleIleAspThrAlaLys 370
 QY 1217 ATGACAGCTTTCGCTCTGCTGCTGAGTACGCTCTACTTCTATGATCCCACTACATA 1276
 Db 371 ValThrSerIleAlaGluAlaThrGlyGlnSerIleTyrPheTyrAspProIleThrAsn 390
 QY 1277 GATCATCCACCAACAGTTTACAGATGTCTTAAAGTTAATGAGACTCCGCGAGATTCGCA 1336
 Db 391 ProGlyThrAlaAlaSerThrAspThrLeuAsnLeuAlaAspAlaAsnSerGlu 410
 QY 1337 CTACAAATATACAGGAACATCATCTTACAGAGAGAAAGTTATCAGACAGAGCGCGCA 1396
 Db 411 IleGluTyrGlyGlyAlaIleValPheSerGlyGlyLysLeuSerProThrGluLysAla 430
 QY 1397 GATTCCTAAAATCTTCTTCCGAAGCTACTACAGCTGTAACTCTTTCAGGAGTACTCTA 1456
 Db 431 IleAlaAlaAsnValThrSerThrIleArgGlnProAlaValLeuAlaArgGlyAspLeu 450
 QY 1457 TCTTTAAACATGGAGTACTCTGAGACTCAGGCACTTCACTCAACAGGAGATTCGTGT 1516

Db	451	valLeuargAspGlyValThrValThrPheLysAspLeuThrGlnSerProGlySerArg	470
Qy	1517	CTCGAAATGGACGTAGGAACCTACTCTAGAACCTCGTGCATATAGC---ACCATAAAACAAT	1573
Db	471	ileLeuMetAspGlyGlyThrThrLeuSerAlaLysGluAlaAsnLeuSerLeuAsnGly	490
Qy	1574	TTGTCTATAATACATCAGTCTTATAGACGGTGCAAGAGGCAAAAAATAGAAACCAAGCT	1633
Db	491	LeuAlaValAsnLeuSerSerLeuAspGlyThrAsnLysAlaLaLeuLysThrGluAla	510
Qy	1634	ACGTCAAAAATCTGACTTATCTTCGGAACCATCACTTATTTGGACCCGACGGGCACGCTTT	1693
Db	511	AlaAspLysAsnIleSerLeuSerGlyThrIleAlaLeuIleAspThrGluGlySerPhe	530
Qy	1694	TATGAAATCATAGTATTAAAGAAATCCCTCAGCTCTACGACATCTTTAGAGCTCMAA----	1747
Db	531	TyrGluAsnHisAsnLeuLysSerAlaSerThrTyrProLeuLeuGluLeuThrAla	550
Qy	1748	---GCTTCTGGAACGTATAACAGCACCGCAGTGTACTCAGATCTATAATATGGGTGAGAAA	1804
Db	551	GlyAlaAsnGlyThrIleLeuGlyAlaLeuSerThrLeuThrLeuGlnGluProGlu	570
Qy	1805	TTCCATTACCGCTATCAGGAACTTGGGCCCAATTGTTGGGGGACAGGGGCTTCTACG	1864
Db	571	ThrHisTyrGlyTyrGlnGlyAsnTrp---GlnLeuSerTrpAlaAsnAlaThrSerSer	589
Qy	1865	---ACTCCAACTTCAACTGACATAAACTGGCTATATTCCTAATCCGACGCGTATCGGC	1921
Db	590	LysIleGlySerIleAsnTrpThrArgThrGlyTyrIleProSerProGluArgLysSer	609
Qy	1922	TCATTAGTCCTTAATAGCTTATGAAATGCATTTATAGATATTAGCTCTCTCCATTATCTT	1981
Db	610	AsnLeuProLeuAsnSerLeuTrpGlyAsnPheIleAspIleArgSerIleAsnGlnLeu	629
Qy	1982	ATGAGAGACTGCAACGAAAGGTTTCAGGGAGACCGTCTTTTGGTGTGCTGATATCT	2041
Db	630	IleGluThrLysSerSerGlyGluProPheGluArgGluLeuTrpLeuSerGlyIleAla	649
Qy	2042	AACTCTCTTCCATTAAGGATAGTACAAAAACACGACGCGGGTTTCGCCATTTTGAGTGGCGGT	2101
Db	650	AsnPhePheTyrArgAspSerMetProThrArgHisGlyPheArgHisIleSerGlyGly	669
Qy	2102	TATGTCATAGAGGAAACCTACATACATCTGTTCAGATAAGATTCTTAGTGTGCTGATTTGT	2161
Db	670	TyrAlaLeuGlyIleThrAlaThrThrProAlaGluAspGlnLeuThrPheAlaPheCys	689
Qy	2162	CAGCTCTTTGGAGAGATAGAGACTACTTTGTAGCTTAAGATCAAGGTACAGTCTACGGA	2221
Db	690	GlnLeuPheAlaArgAspArgAsnHisIleThrGlyLysAsnHisGlyAspThrTyrGly	709
Qy	2222	GGAATCTCTATTACCAAGCAACGAAACCTATATCTCTCTCTTCCATCAACTA-----	2275
Db	710	AlaSerLeuTyrPheHisHisThrGluGlyLeuPheAspIleAlaAsnPheLeuTrpGly	729
Qy	2276	-----CGGCTCTGTTCTGTTGCTTATGTTCTCTACAGAGATTCCTGTCTCTCTTT	2323
Db	730	LysAlaThrArgAlaProTrpValLeuSerGluIleSerGlnIleIleProLeuSerPhe	749
Qy	2324	TCAGGAAACCTTAGCTACACCCATACGGATAACGATCTGAAACCAAGTATACACATAT	2383
Db	750	AspAlaLysPheSerTyrLeuHisThrAspAsnHisMetLysThrTyrTyrThrAspAsn	769
Qy	2384	CCTACTCTTAAAGGAAGCTGGGGAATGATAGTTTCGTTTAAAGATTCGGGTGGAAGCT	2443
Db	770	SerIleIleLysGlySerTrpArgAsnAspAlaPheCysAlaAspLeuGlyAlaSerLeu	789
Qy	2444	CCGATTTCCTTAGTAAGAGTCTCTATTTCAGCAGTACATGCCCTTCATGAATATGCAG	2503
Db	790	ProPheValIleSerValProTyrLeuLeuLysGluValGluProPheValLysValGln	809
Qy	2504	TTTGTCTATCCATCAGGAAGGTTTAAAGACAGGGAAACAGAGCTCTGTGAATTGGA	2563

431	IleAlaAlaAsnValThrSerThrIleArgGlnProAlaValLeuAlaArgGlyAspLeu	450
1457	TCCTTTAAACATGAGTGACTCTGCAGACTCAGGCATTCCTCAACAGCGAGATCTCGT	1516
451	ValLeuArgAspGlyValThrValThrPheLysAspLeuThrGlnSerProGlySerArg	470
1517	CTCGAAATGGACGTAGGAACACTCTCTAGAACCTGCTGTACTAGC--ACCATAAACAAT	1573
471	IleLeuMetaspGlyGlyThrThrLeuSerAlaLysGluAlaAsnLeuSerLeuAsnGly	490
1574	TTGTGTCATTACATCAGTCTTATAGACGGTGCAGAAAGCGCAAAATATAGAAACCAAGCT	1633
491	LeuAlaValAsnLeuSerLeuAspGlyThrAsnLysAlaAlaLeuLysThrGluAla	510
1634	AGCTCAAAAATCTGACTTTATCTGGAACCATCAGTCTTATTTGGACCGACGGCAGCTTT	1693
511	AlaAspLysAsnIleSerLeuSerGlyThrIleAlaLeuIleAspThrGluGlySerPhe	530
1694	TATGAAATCATAGTTTAAAGAAATCTCTAGTCCTACGATCTCTAGAGCTCAAA-----	1747
531	TyrGluAsnHisAsnLeuLysSerAlaSerThrTyrProLeuLeuGluLeuThrAla	550
1748	---GCTTCTGGAATGTAAACAGACCGCAGGTGACTCCAGATCTCTATTAATGGGTGAGAAA	1804
551	GlyAlaAsnGlyThrIleThrLeuGlyAlaLeuSerThrLeuThrLeuGlnGluProGlu	570
1805	TTCCATTACGCTATCAGGAACTTGGGGCCCAATTTGTTGGGACAGAGGGCTCTCTACG	1864
571	ThrHisTyrGlyTyrGlnGlyAsnTrp---GlnLeuSerTrpAlaAsnAlaThrSerSer	589
1865	---ACTGCAACCTTCAACTCGACTAAACCTGGCTATATTCCTTAATCCGACGGTATCGGC	1921
590	LysIleGlySerIleAsnTrpThrArgThrGlyTyrIleProSerProGluArgLysSer	609
1922	TCCTTAGTCCCTATAGCTTATGGAATGCATTTATAGATATTAGCTCTCTCCATATCTT	1981
610	AsnLeuProLeuAsnSerLeuTrpGlyAsnPheIleAspIleArgSerIleAsnGlnLeu	629
1982	ATGAGAGCTGCAGAAAGGGTTGACGGAGACCGCTGCTTTTGGTGTGCTGGATTATCT	2041
630	IleGluThrLysSerSerGlyGluProPheGluArgGluLeuTrpLeuSerGlyIleAla	649
2042	ACTCTTCTCCATAGGATAGTACAAAACACGACGGGGTTTCGCCATTGAGTGGCGGT	2101
650	AsnPhePheTyrArgAspSerMetProThrArgHisGlyPheArgHisIleSerGlyGly	669
2102	TATGTCATAGAGGAAACACTACATCTGTTCAGATAAGATTCCTAGTCTGCTGCAATTTGT	2161
670	TyrAlaLeuGlyIleThrAlaThrProAlaGluAspGlnLeuThrPheAlaPheCys	689
2162	CAGCTCTTTGGAAGAGATAGAGACTACTTTGTAGCTTAAAGAACAGGTACAGCTACCGA	2221
690	GlnLeuPheAlaArgAspArgAsnHisIleThrGlyLysAsnHisGlyAspThrTyrGly	709
2222	GGAACTCTCTATTACACGACCAAGAAACCTATATCTCTCTCTTGGCAAACTA-----	2275
710	AlaSerLeuTyrPheHisHisThrGluGlyLeuPheAspIleAlaAsnPheLeuTrpGly	729
2276	-----CGCCCTTGTTGCTGCTCTATGTTCTTACAGAGATTCCTGTTCTCTTTT	2323
730	LysAlaThrArgAlaProTrpValLeuSerGluIleSerGlnIleIleProLeuSerPhe	749
2324	TCAGGAAACCTTAGCTACACCCATACGGATAACGATCTGAAACCAACGATATACACATAT	2383
750	AspAlaLysPheSerTyrLeuHisThrAspAsnHisMetLysThrTyrTyrThrAspAsn	769
2384	CCTACTGTTTAAAGGAGCTGGGGGAATGATAGTTTTCGCTTTAGATTCGGTGGAGAGCT	2443
770	SerIleIleLysGlySerTrpArgAsnAspAlaPheCysAlaAspLeuGlyAlaSerLeu	789
2444	CGAATTGCTTAGATGAAGTCTCTATTGACAGCATACATGCCCTTCTAGAAATTCGAC	2503
790	ProPheValIleSerValProTyrLeuLeuGlyValGluProPheValLysValGln	809

Db 411 IleGluTyrGlyGlyAlaIleValPheSerGlyGluLysLeuSerProThrGluLysAla 430
 QY 1397 GATTCTAAATCTTCTGAGAGCTACTACAGCTGTACTCTTTCAGAGGTACTCTA 1456
 Db 431 IleAlaAlaSerValThrSerThrIleArgGlnProAlaValLeuAlaArgGlyAspLeu 450
 QY 1457 TCTTTAAACATGAGTGACTCTCGACACTCAGGCATTCCTCAACAGGCGAGATTCTCGT 1516
 Db 451 ValLeuArgAspGlyValThrValThrPheLysAspLeuThrGlnSerProGlySerArg 470
 QY 1517 CTCGAAATGGAGCTAGGACTCTCTAGAACCTCTGATAGCTAGC---ACCATAACAAT 1573
 Db 471 IleLeuMetAspGlyGlyThrThrLeuSerAlaLysGluAlaAsnLeuSerLeuAsnGly 490
 QY 1574 TTGTCATTACATCAGTCTCTATAGACGGTGCAGGCAAGGCAAAATAGAAACCAAGCT 1633
 Db 491 LeuAlaValAsnLeuSerLeuAspGlyThrAsnLysAlaAlaLeuLysThrGluAla 510
 QY 1634 ACGTCAAAATCTGACTTTATCGGAACCATCTTTATTTGGACCGGACGGCAGCTTT 1693
 Db 511 AlaAspLysAsnIleSerLeuSerGlyThrIleAlaLeuIleAspThrGluGlySerPhe 530
 QY 1694 TATGAATATCATCTTAAGAACTCTCAGTCTCAGCATCTCTAGAGCTCAA----- 1747
 Db 531 TyrGluAsnHisAsnLeuLysSerAlaSerThrTyrProLeuLeuGluLeuThrThrAla 550
 QY 1748 ---GCTTCTGGAATCTGAACAGCACCGCAGTCTCAGATCCTATAATGGGTGAGAA 1804
 Db 551 GlyAlaAsnGlyThrIleThrLeuGlyAlaLeuSerThrLeuThrLeuGluProGlu 570
 QY 1805 TTCATTACGGCTATCAGGAGACTGGGGCCCAATTTGTTGGGGACAGGGCTTCTAGC 1864
 Db 571 ThrHisTyrGlyTyrGlnGlyAsnTrp---GlnLeuSerTrpAlaAsnAlaThrSerSer 589
 QY 1865 ---ACTGCAACCTTCAACTGACTAAACTGGCTATATCTCAATCCGAGCGTATCGGC 1921
 Db 590 LysIleGlySerIleAsnTrpThrArgThrGlyTyrIleProSerProGluAlaGlySer 609
 QY 1922 TCTTTAGTCCCTATAGCTTATGGATGCAATTAATAGATTAATAGCTCTCTCCATATCTT 1981
 Db 610 AsnLeuProLeuAsnSerLeuTrpGlyAsnPheIleAspIleArgSerIleAsnGlnLeu 629
 QY 1982 ATGGAGACTGCAACGAAAGGTTGTCAGGAGACCGTCTGTTTGGTGTGCTGATATCT 2041
 Db 630 IleGluThrLysSerSerGlyGluProPheGluArgGluLeuTrpLeuSerGlyIleAla 649
 QY 2042 AACTTCTTCCATAGAGATAGAGACTACTTGTAGCTAAGATCAAGTACAGTCTACGGA 2101
 Db 650 AsnPhePheTyrArgAspSerMetProThrArgHisGlyPheArgHisIleSerGlyGly 669
 QY 2102 TATGTCATAGGAGAACTCATATCTTGTTCAGATAAGATCTTGTAGTCTGCTGATTTGT 2161
 Db 670 TyrAlaLeuGlyIleThrAlaThrThrProAlaGluAspGlnLeuThrPheAlaPheCys 689
 QY 2162 CAGCTCTTTGGAGAGATAGAGACTACTTGTAGCTAAGATCAAGTACAGTCTACGGA 2221
 Db 690 GlnLeuPheAlaArgAspArgAsnHisIleThrGlyLysAsnHisGlyAspThrTyrGly 709
 QY 2222 GGAATCTCTATTACGACACACAGCAACCTATATCTCTCTCTCTGCAACTA----- 2275
 Db 710 AlaSerLeuTyrPheHisThrGluGlyLeuPheAspIleAlaAsnPheLeuTrpGly 729
 QY 2276 -----CGGCTTGTGTTGCTTGTCTTATGTTCTCAGAGATTCCTGTTCTCTTT 2323
 Db 730 LysAlaThrArgAlaProTrpValLeuSerGluIleSerGlnIleProLeuSerPhe 749
 QY 2324 TCAGGAACCTTAGCTACCCATACGATACGATACGATCGAAGCAACAGTATACACATAT 2383
 Db 750 AspAlaLysPheSerTyrLeuHisThrAspAsnHisMetLysThrTyrThrAspAsn 769
 QY 2384 CCTACTGTTTAAAGGAGCTGGGGGATGATAGTTTCGCTTTAGAAATTCGGTGGGAAGCT 2443
 Db 770 SerIleIleLysGlySerTrpArgAsnAspAlaPheCysAlaAspLeuGlyAlaSerLeu 789

QY 2444 CCGATTTCCTAGATGAAGTCTCTATTGACAGCTACATGCCCTTCATGAATTCGAC 2503
 Db 790 ProPheValIleSerValProTyrLeuLeuLysGluValGluProPheValLysValGln 809
 QY 2504 TTTGCTCTATCAGCATCAGGAGGTTTAAAGAACAGGGAACAGAGCTCGTGAATTTGA 2563
 Db 810 TyrIleTyrAlaHisGlnAspPheTyrGluArgTyrAlaGluArgAlaPheAsn 829
 QY 2564 AGTAGCGCTCTGTAATCTTGCCTTACCTATCCGATCCGATTCGATTAAGGAATCAGAC 2623
 Db 830 LysSerGluLeuIleAsnValGluIleProIleGlyValThrPheGluArgAspSerLys 849
 QY 2624 TGCCAAGATGCAACGTAACATCTAACTCTTGGTATATCTGTGATCTTCTGTAGTAAC 2683
 Db 850 SerGluLysGlyThrTyrAspLeuThrLeuMetTyrIleLeuAspAlaTyrArgAsn 869
 QY 2684 CCCGACTGTACGACCAACACTGCGAATGAGGTGATCTCTGGAAACCTTCGTTACGAT 2743
 Db 870 ProLysCysGlnThrSerLeuIleAlaSerAspAlaAsnTrpMetAlaTyrGlyThrAsn 889
 QY 2744 TTGGCAAGACAAAGCTTTAGTCTCTGTCAGGGAACCAATTTTCTTAACTCAAATTT 2803
 Db 890 LeuAlaArgGlnGlyPheSerValArgAlaAlaAsnHisPheGlnValAsnProHisMet 909
 QY 2804 GAAGCTTTAGCCAAATTTCTTTTGAATTCGCGGTGATCTCGCAATTCAGATAGAC 2863
 Db 910 GluIlePheGlyGlnPheAlaPheGluValArgSerSerArgAsnTyrAsnThrAsn 929
 QY 2864 TTAGAGCAAAATACCAATTC 2884
 Db 930 LeuGlySerLysPheCysPhe 936
 RESULT 9
 D86546
 C:Species: Chlamydomonas reinhardtii [imported] - Chlamydomonas reinhardtii
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 15-Jun-2001
 C:Accession: D86546
 R:Shirai, M.; Hirakawa, H.; Kinoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;
 Nucleic Acids Res. 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of Chlamydomonas reinhardtii
 A:Reference number: AB6491; MUID:20330349; PMID:10871362
 A:Accession: D86546
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-930 <STO>
 A:Cross-references: GB:BA000008; NID:g978818; PIDN:BA098654.1; GSPDB:GN00142
 A:Experimental source: strain J138
 C:Genetics:
 C:Superfamily: Chlamydomonas reinhardtii polypeptide outer membrane protein G
 Alignment Scores:
 Pred. No.: 4.8e-115 Length: 930
 Score: 1940.00 Matches: 416
 Percent Similarity: 60.97% Conservative: 165
 Best Local Similarity: 43.65% Mismatches: 324
 Query Match: 35.87% Indels: 48
 DB: 2 Gaps: 17
 US-09-428-122-1 (1-3000) x D86546 (1-930)
 QY 101 ATGAAGTCTCTCTTCCCAAGTTGTTATTTTCTACATTTGCTATTTTCCCTTTCTATG 160
 Db 1 MetLysIleProLeuHisLysLeuLeuIleSerSerThrLeuValThrProIleLeu 20
 QY 161 ---ATTGCTACC-----GAGACAGTTTGGATTCAGTGCAGTTCGATGGGAAT 208
 Db 21 SerIleAlaThrTyrGlyAlaAspAlaSerLeuSerProThrAspSerPheAspGlyAla 40
 QY 209 AAAAATGGTAATTTTTCAGTTTCGTCAGAGTTCAGGAGTTCAGGAGTTCAGTCTATTT 268

Db 41 GlyGlySerThrPheThrProLysSerThrAlaAspAlaAsnGlyThrAsnTyrValLeu 60
QY 269 AAGGAAATGCTACTCTAGAAAAATATCTCTGGAACAGCAGCAATCACAAAAAGCTGT 328
Db 61 SerGlyAsnValTyrIle---AsnAspAlaGlyLysGlyThrAlaLeuThrGlyCysCys 79
QY 329 TTTAAACAACACTAAGGGCGATTGCTACTTTCACAGTAACGGGAACCTCTATTGTTCCAA 388
Db 80 PheThrGluThrThrGlyAspLeuThrPheThrGlyLysGlyTyrSerPheSerPheAsn 99
QY 389 ACGGTGGATGACAGGACTGTAGAGGGCTGCTTTAACACAGCAGCGTGGTAGATAATCT 448
Db 100 ThrValAspAlaGlySerAsnAlaGlyAlaAla---SerThrThrAlaAspLysAla 118
QY 449 ACCAGCTTTATAGGGTTTCTTCGGTATCTTTTATGCTCTCTGGAAGTTCGATAACT 508
Db 119 LeuThrPheThrGlyPheSerAsnLeuSerPheIleAlaAlaProGlyThrThrValAla 138
QY 509 ACCGGCAAGGAGCCGTTAGCTGTCTACGGGTAGCTTGAGTTGACAAAATGTCAGT 568
Db 139 SerGlyLysSerThrLeuSer---SerAlaGlyAlaLeuAsnLeuThrAspAsnGlyThr 157
QY 569 TTGCTCTTCAGCAAAAATTTCACCGAT-----AATGGGGTGTCTATCACCGCA 619
Db 158 IleLeuPheSerGlnAsnValSerAsnGlnAlaAsnAsnGlyGlyAlaIleThrAla 177
QY 620 AAAACTCTTTCAATTAACAGGAGCTACAATGTGCTGTGTTTCTGAAAATPACCTCTCA 679
Db 178 LysThrLeuSerIleSerGlyAsnThrSerSerIleThrPheThrSerAsnSerAlaLys 197
QY 680 AAGAAGGGGAGCCATTCAGCTCCGATGCTCCGATGCTCCGATGCTCCGATGCTCCGATGCT 739
Db 198 LysLeuGlyGlyAlaIleTyrSerSerAlaAlaAlaSerIleSerGlyAsnThrGlyGln 217
QY 740 GTCCTCTTTCTGCAATCTCTTCGGATCTCTGAGCTGCAATTTTACAGAGCTCG 799
Db 218 LeuValPheMetAsnAsnLysGlyGluThrGlyGlyAlaLeuGlyPheGluAlaSer 237
QY 800 GTGACTATTCTTAATAGTCTTAAGTTCTTCTTATGCAATAAGTCCAGGAGCGAGC 859
Db 238 SerSerIleThrGlnAsnSerSerLeuPhePheSerGlyAsnThrAlaThrAspAlaAla 257
QY 860 TCCTCAACAACGGGGATATGTCAGAGGTGCTATCTGCTTATAAAGTACAGAT 919
Db 258 Gly-----LysGlyGlyAlaIleTyrCysGluLysThrGlyGluThr 271
QY 920 ACTAAGGTCCACTGGAATCAGATGTTTACTCTTCAGCAACAATACATCGACACA 979
Db 272 ProThrLeuThrIleSerGlyAsnLysSerLeuThrPheAlaGluAsnSerSerValThr 291
QY 980 GCGGGAGGACTATCTATGTGAAAAGCTCGAAGCTGCTTCGGAGGACTTACCTATTC 1039
Db 292 GlnGlyGlyAlaIleCysAlaHisGlyLeuAspLeuSerAlaAlaGlyProThrLeuPhe 311
QY 1040 AGTAGAAATAGTGTCAATGGAGTACAGCTCCCTAAAGGTGGAGCCATAGCTATCCAGAT 1099
Db 312 SerAsnAsnArgCysGlyAsnThrAlaAlaGlyLysGlyAlaIleAlaIleAlaAsp 331
QY 1100 AGTGGGAATAGATTATCCCGCATAGTGTGACATGCTTTTATGGGAATACATC 1159
Db 332 SerGlySerLeuSerLeuSerAlaAsnGlnGlyAspIleThrPheLeuGlyAsnThrLeu 351
QY 1160 ACTTCTACTACT---CCTGGAGCAATAGAAGTAGTATCGATCTAGGACGAGTCCAAAG 1216
Db 352 ThrSerThrSerAlaProThrSerThrArgAsnAlaIleTyrLeuGlySerSerAlaLys 371
QY 1217 ATCAGAGCTTTGCTGCTGCTGAGAGCCATCTACTTCTATGATCCCATAT 1270
Db 372 IleThrAsnLeuArgAlaAlaGlnGlyGlnSerIleTyrPheTyrAspProIleAlaSer 391
QY 1271 ---ACTACAGGATCATCCACAACAGTATGCTTTAAAAGTTAATGAGACTCCGCA 1327
Db 392 AsnThrThrGlyAlaSer-----AspValLeuThrIleAsnGlnProAspSer 407

QY 1328 GATTCTGCACACTACAATATACAGGGAACATCATCTTACAGGAGAAAAGTTATCAGAGCA 1387
Db 408 AsnSerProLeuAspTyrSerGlyThrIleValPheSerGlyGlyLysLeuSerAlaAsp 427
QY 1388 GAGGCGGCGAGATTCTTAAATACTTACTTCGAAGCTACTACAGCCTGTAACTCTTTTCAGGA 1447
Db 428 GluAlaLysAlaAlaAspAsnPheThrSerIleLeuLysGlnProLeuAlaLeuAlaSer 447
QY 1448 GGTACTCTATCTTTTAAACATCGAGTACTCTGAGACTCAGGCATCTCACTACACAGCA 1507
Db 448 GlyThrLeuAlaLeuLysGlyAsnValGluLeuAspValAsnGlyPheThrGlnThrGlu 467
QY 1508 GATTCTCTGCTCGAATAGAGCTAGGAATCTACTCTAGAACCTGTCTGATCTAGCACCATA 1567
Db 468 GlySerThrLeuLeuMetGlnProGlyThrLysLeuLys---AlaAspThrGluAlaIle 486
QY 1568 AAC-----AATTTGTCATTAAACATCATCATGTTCTATAGCGGTGCAAGAGGCAAAATA 1621
Db 487 SerLeuThrLysLeuValValAspLeuSerAlaLeuGluGlyAsnLysSerValSerIle 506
QY 1622 GAACCAAGACTACGCTCAAAAATCTGACTTTTATCTGGAACCATCTTTATTTCGACCG 1681
Db 507 GluThrAlaGlyAlaAsnLysThrIleThrLeuThrSerProLeuValPheGlnAspSer 526
QY 1682 ACGGCGACCTTTTATGAAATCATAGTTTAAAGAAATCTCTCAGTCTCAGCATCTTAGAG 1741
Db 527 SerGlyAsnPheTyrGluSerHisThrIleAsnGlnAlaPheThrGlnProLeuValVal 546
QY 1742 CTCAAAGCTCTC-----GGAACTGTAAACAGCACCGCA 1774
Db 547 PheThrAlaAlaThrAlaAlaSerAspIleTyrIleAspAlaLeuLeuThrSerProVal 566
QY 1775 GTGACTCCAGATCTCTATATGCTGAGAAATTCATTCAGGCTATCAGGGAACCTTGGGCG 1834
Db 567 GlnThrProGluPro-----HisTyrGlyTyrGlnGlyHisTrpGlu 580
QY 1835 CCAATTTGTTGGGGAGCGGCTCTACGACTGCACTTCACTGAGCACTTAACTGCACTGCGC 1894
Db 581 AlaThrTrpAlaAspThrSerThrAlaLysSerGlyThrMetThrTrpValThrThrGly 600
QY 1895 TATATTCTTAATCCCGAGGCTATCGGCTCTTTAGTCCCTTAATAGCTTATGGAATCATTT 1954
Db 601 TyrAsnProAsnProGluArgAlaSerValValProAspSerLeuTrpAlaSerPhe 620
QY 1955 ATGATATTAGTCTCTCCTATCTTATGAGAGCTGCAACCAAGGTTGCGAGGAGAC 2014
Db 621 ThrAspIleArgThrLeuGlnGlnIleMetThrSerGlnAlaAsnSerIleTyrGlnGln 640
QY 2015 CGTGTCTTTTGTGTCTGATTTATCTACTTCTTCCATAGATAGTACAAAACACGA 2074
Db 641 ArgGlyLeuTrpAlaSerGlyThrAlaAsnPhePheHisLysAspLysSerGlyThrAsn 660
QY 2075 CCGGGTTTCGCCATTTGAGTGGCGGTATGTCATAGGAGGAACCTCATACTTGTTC 2134
Db 661 GlnAlaPheArgHisLysSerTyrGlyTyrIleValGlyGlySerAlaGluAspPheSer 680
QY 2135 GATAAGATCTTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2194
Db 681 GluAsnIlePheSerValAlaPheCysGlnLeuPheGlyLysAspLysPheIle 700
QY 2195 GCTAAGAATCAAGGTACAGTCTCAGGAGGAACCTCTCTATTACCAACCAACCAACCTAT 2254
Db 701 ValGluAsnThrSerHisAsnTyrLeuAlaSerLeuTyrLeuGlnHisArgAlaPheLeu 720
QY 2255 ATCTCTCTCTCTGCAAACTACGCGCTTGT-----TCCTGCTCTTATGTTCTCTACAGAG 2308
Db 721 GlyGlyLeuPro-----MetProSerPheGlySerIleThrAspMetLeuLysAsp 737
QY 2309 ATTCTGTTCTCTTTTTCAGGAACCTTAGCTACACCCATACGATAACGATCTGAAACCC 2368
Db 738 IleProLeuIleLeuAsnAlaGlnLeuSerTyrSerTyrThrLysAsnAspMetAspThr 757

QY	1271	--	ACTACGAGTATCCACAACAGTTTACAGATGCTCTAAAGATTTATGAGACTCCGGCA	1327
Db				
QY	1328	GATTCTGCATCAATATACAGGAGCATCATCTTCACAGAGCAAGAAAGTTATCAGAGACA	1387	
Db				
QY	1388	GAGCGCGCAGATTCTAAAAATCTTACTTCGAAGCTACTACAGCCTCTTAACCTTTTCAGGA	1447	
Db				
QY	1448	GGTACTCTATCTTAAACATGAGGAGTACTCTCCAGACTCAGGCATTTCACCTCAACAGGCA	1507	
Db				
QY	1508	GATTCTCGTCTCGAAATGGAGTAGGAACCTACTCTAGAACCTGCTGATACGACCATTA	1567	
Db				
QY	1568	AAC-----AATTGGTCATTAAACATCAGTTCCTATAGACGGTCCAAAGAGGCCAAATA	1621	
Db				
QY	1622	GAACCAAAAGCTACGTCAAAAATCTGACTTTATCTGAAACCATCACTTTATTTGGACCGG	1681	
Db				
QY	1682	ACGGGACGTTTATGAAATCATAGTTTAAAGAAATCTCAGTCTCTACGACATCTTAGAG	1741	
Db				
QY	1742	CTCAAGACTTCT-----GGAACTGTAAACAAGCACCGCA	1774	
Db				
QY	1775	GTGCTCCAGATCCTAATGGGTGAGAAATCCATTACGGCTATCAGGAACTTGGGC	1834	
Db				
QY	1835	CCAATTGTTGGGGACAGGGGCTTCTACAGCTGCACCTTCAACTGGACTAAACTGGC	1894	
Db				
QY	1895	TATATTCTTAATCCCGAGCGTATCGGCTCTTTAGTCCCTAATAGTCTTGAATGCATTT	1954	
Db				
QY	1955	ATAGATATTAGCTCTCTCCATTATCTTATGGAGACTGCAACCTTCAACTGGACTAAACTGGC	2014	
Db				
QY	2015	CGTCTTTTGGTGTGCTGGATATCTAACTTCTTCATAGAGTACAGTACATAAAACACGA	2074	
Db				
QY	2075	CGCGGTTTCCCATTTGAGTGGCGGTATTGTCATAGGAGAACTTACATCTTGTTC	2134	
Db				
QY	2135	GATTAAGATTCTTAGTGTGCTGCTTTTGTACAGTCTTTTGGAGAGATAGAGACTTCTTGT	2194	
Db				
QY	2195	GCTAAGAAATCAAGGTACAGTCTACGGAGGAACTCTTATTACCAGCACCAACGAAACCTAT	2254	
Db				
QY	2255	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	2308	
Db				
QY	2308	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	2372	
Db				
QY	2372	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	2436	
Db				
QY	2436	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	2500	
Db				
QY	2500	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	2564	
Db				
QY	2564	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	2628	
Db				
QY	2628	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	2692	
Db				
QY	2692	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	2756	
Db				
QY	2756	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	2820	
Db				
QY	2820	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	2884	
Db				
QY	2884	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	2948	
Db				
QY	2948	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	3012	
Db				
QY	3012	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	3076	
Db				
QY	3076	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	3140	
Db				
QY	3140	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	3204	
Db				
QY	3204	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	3268	
Db				
QY	3268	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	3332	
Db				
QY	3332	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	3396	
Db				
QY	3396	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	3460	
Db				
QY	3460	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	3524	
Db				
QY	3524	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	3588	
Db				
QY	3588	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	3652	
Db				
QY	3652	ATCTCTCTCTCTCAAACTACGGCCTTGT-----TCGTTGTCTTATGTTCTCTACAGAG	3716	
Db				

QY	2309	ATTCTGTGTTCTCTTTTCAGAAACCTTAGCTACACCCATACGGATACCGATCTGGAACACC	2368
Db	738	ileProLeuileLeuAsnAlaGlnLeuSerTyrSerTyrThrLysAsnAspMetAspThr	757
QY	2369	AAGTATACAAATATCTACTCTGTAAAGGAAGCTGGGGAATCATAGTTTTCGCTTTAGAA	2428
Db	758	ArgTyrThrSerTyrProGluAlaGlnGlySerTyrThrAsnAsnSerGlyAlaLeuGlu	777
QY	2429	TTTCGTTGGAAAGACTCCGATTTCCTTA---GATGAAAGTGCCTCTATTGTAGCAGTACATG	2485
Db	778	LeuGlyGlySerLeuAlaLeuTyrLeuProLysGluAlaProPheGlnGlyTyrPhe	797
QY	2486	CCCTTCATGAAATTGCAGTTTGTCTATGCACATCAGGAAGTTTAAAGAACAGGGAACA	2545
Db	798	ProPheLeuIysPheGlnAlaValTyrSerArgGlnGlnAsnPhenylSerGlyAla	817
QY	2546	GAAGCTCGTGAATTTGGAAGTAGCCGCTCTGTGTAATCTTGCCTTACCTATCGGGATCCGA	2605
Db	818	GluAlaArgAlaPheAspGlyAspLeuValAsnCysSerIleProValGlyIleArg	837
QY	2606	TTTGATAAGGAATCAGACTGCCAAGTAGCAACCTCAACTAACTCTCGTTATATCTGTG	2665
Db	838	LeuGluIysIleSerGluAspGluLysAsnAsnPhelIleSerLeuAlaIyrIleGly	857
QY	2666	GATCTTGTGTTAGTAAACCCCGACTGTACGACAACTGCGAATTAGCGGTGATTTCTTG	2725
Db	858	AspValTyrArgIysAsnProArgSerArgThrSerLeuMetValSerGlyAlaSerTyr	877
QY	2726	AAACCTTCGGTAGCAATTTGGCAAGCAGCTTTAGTCCCTCGTCGACGGAACCACTTTT	2785
Db	878	ThrSerLeuCysLysAsnLeuAlaArgGlnAlaPheLeuAlaSerAlaGlySerHisLeu	897
QY	2786	TGCTTTAACTCAAAATTTGAAGCCCTTTAGCCAAATTTCTTTTGAATTCGTGGTCACTCT	2845
Db	898	ThrLeuSerProHisValGluLeuSerGlyGluAlaAlaTyrGluLeuArgGlySerAla	917
QY	2846	CGCAATTACATGTAGACTTAGGACGAAATACCAATTC	2884
Db	918	HisIleTyrAsnValAspCysGlyLeuArgTyrSerPhe	930

RESULT 11

D72078

polymorphic outer membrane protein g family - Chlamydophila pneumoniae (strain CWL025)

C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 15-Jun-2001

C;Accession: D72078

R;Kalmann, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, Nature Genet. 21, 385-389, 1999

A;Title: Comparative Genomes of Chlamydia pneumoniae and C. trachomatis.

A;Reference number: A72000; MUID:99206606; PMID:10192388

A;Accession: D72078

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-930 <ARN>

A;Cross-references: GB:AE001627; GB:AE001363; NID:g4376721; PIDN:AAD18590.1; PID:g4376721

A;Experimental source: strain CWL025

Db 721 GlyGlyLeuPro-----MetProSerPheGlySerIleThrAspMetLeuLysAsp 737
QY 2309 ATTCTGTTCTCTTTTCAGGAAACCTTAGCTACACCCATACCGGATACAGTCTGAAACC 2368
Db 738 IleProLeuIleLeuAsnAlaGlnLeuSerTyrSerTyrThrLysAsnAspMetAspThr 757
QY 2369 AAGTATACAAATATCTCTACTGTTTAAAGGAGCTGGGGAATGATAGTTTCGCTTTAGAA 2428
Db 758 ArgTyrThrSerTyrProGluAlaGlnGlySerTyrThrAsnAsnSerGlyAlaLeuGlu 777
QY 2429 TTCGGTGAAGAGCTCCGATTTGCTTAA---GATGAAAGTGTCTATTTGAGCAGTACATG 2485
Db 778 LeuGlyGlySerLeuAlaLeuTyrLeuProLysGluAlaProPhePheGlnGlyTyrPhe 797
QY 2486 CCCTTCATGAATTCGATTTGCTATGACATCAGGAAGTTTAAAGAACAGGGAACA 2545
Db 798 ProPheLeuLysPheGlnAlaValTyrSerArgGlnGlnAsnPheLysGluSerGlyAla 817
QY 2546 GAAGCTCGTGAATTTGGAAGTAGCGCTTGTGTAATCTTGCTTACCTATCGGATCCGA 2605
Db 818 GluAlaArgAlaPheAspAspGlyAspLeuValAsnCysSerIleProValGlyIleArg 837
QY 2606 TTGATGAAGAAATCAGACTGCCAAGTCCAACTGACAACTTAATCTTGTGTTACTGTG 2665
Db 838 LeuGluLysIleSerGluAspGluLysAsnAsnPheGluIleSerLeuAlaTyrIleGly 857
QY 2666 GATCTTGTTCGTAGTACCCGACTGACGACAACTGCGAATACCGGTGATCTTTGG 2725
Db 858 AspValTyrArgLysAsnProArgSerArgThrSerLeuMetValSerGlyAlaSerTyr 877
QY 2726 AAAACCTTCGTGACGAATTTGGCAAGCAAGCTTTAGTCTCTGTCGAGGAAACCAATTT 2785
Db 878 ThrSerLeuCysLysAsnLeuAlaArgGlnAlaPheLeuAlaSerAlaGlySerHisLeu 897
QY 2786 TGCTTTAACTCAATTTTGAAGCTTTAGCCAAATTTCTTTGAAATTCGCTGGGTCTATC 2845
Db 898 ThrLeuSerProHisValGluLeuSerGlyGluAlaAlaTyrGluLeuArgGlySerAla 917
QY 2846 CGCAATTTACAACTTAGCTTAGGAGCAAAATACCAATTC 2884
Db 918 HisIleTyrAsnValAspCysGlyLeuArgTyrSerPhe 930
RESULT 12
F81591
polymorphic membrane protein G family CP0302 [imported] - Chlamydothila pneumoniae (stra
C;Species: Chlamydothila pneumoniae, Chlamydia pneumoniae
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 15-Jun-2001
C;Accession: F81591
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150295; PMID:10684935
A;Accession: F81591
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-949 <REA>
A;Cross-references: GB:AE002161; NID:g7189226; PIDN:AAF38159.1; PID:g718922
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: CP0302
C;Superfamily: Chlamydothila pneumoniae polymorphic outer membrane protein G
Alignment Scores:
Pred. No.: 1,73e-110 Length: 949
Score: 1868.00 Matches: 412
Percent Similarity: 60.49% Conservative: 176
Best Local Similarity: 42.39% Mismatches: 334
Query Match: 34.54% Indels: 50
DB: 2 Gaps: 22
US-09-428-122-1 (1-3000) x F81591 (1-949)

QY 50 TTGGTGTATCTCTTAAATATTAATTAATTT-----CAAAATCAAGTATA 91
Db 1 MetIleTyrLeuPheCysPheTyrIleAspAlaAsnSerSerLeuLysAsnLysSerIle 20
QY 92 TATTTTCAATGAAGTCTTCTTCCCAAGTTTGTATTTTCTACATTTGCTATTTCCCT 151
Db 21 -----ThrMetLysThrSerIleProTyrValLeuValSerSerValLeuAlaPheSer 38
QY 152 -----TTGCTCTATGATTGTCACGAGACAGTTTTGGATTCAAGTGCAGTTCGATGGG 205
Db 39 CysHisLeuGlnSerLeuAlaAsnGluLeuLeuSerProAspSerPheAsnGly 58
QY 206 AAT-----AAAATAGTAAATTTTTCAGTTCGTGAGAGTCAAGAGATGCTGAACTACTAC 262
Db 59 AsnIleAspSerGlyThrPheThrProLysThrSer-----AlaThrThrTyr 74
QY 263 CTATTTAAGGGAATGCTCACTCTAGAAAATATTCCTGGAACAGGACAGCAATCACAAA 322
Db 75 SerLeuThrGlyAspValPhePheTyrGlu---ProGlyLysGlyThrProLeuSerAsp 93
QY 323 AGCTGTTTTTAAACACACTAAGGCGGATTTGACTTTTCACAGGTAACGGGAACCTCTATTG 382
Db 94 SerCysPheLysGlnThrThrAspAsnLeuThrPheLeuGlyAsnGlyHisSerLeuThr 113
QY 383 TTCCAAACCGTGTGATGACGAGCTGTAGCAGGGCTGCTTTAAACAGCAGCGTGTAGAT 442
Db 114 PheGlyPheIleAspAlaGlyThrHisAlaGlyAlaAlaAala---SerThrThrAlaAsn 132
QY 443 AAATCTACACAGTTTATAGGTTTCTTCTGCTATCTTTTATTCGCTCTCTGGAATTCG 502
Db 133 LysAsnLeuThrPheSerGlyPheSerLeuLeuSerPheAspSerSerProSerThrThr 152
QY 503 ATAACCTACCGCAAGAGCGCTTAGCTCTCTACGGGTAGTGTAGTTTGCACAAAAT 562
Db 153 ValThrThrGlyGlnGlyThrLeuSer---SerAlaGlyGlyValAsnLeuGluAsnIle 171
QY 563 GTCAGTTTCTCTTACGAAAACTTTTCAAGGATAATGCGCGTCTATACCGCAAAA 622
Db 172 ArgLysLeuValValAlaGlyAsnPheSerThrAlaAspGlyGlyAlaIleLysGlyAla 191
QY 623 ACTCTTTTCATTAACAGGACTACAATGTCAGTCTGTTTCTTCAAAATACCTCTCAAAG 582
Db 192 SerPheLeuLeuThrGlyThrSerGlyAspAlaLeuPheSerAsnAsnSerSerSerThr 211
QY 683 AAAGCGGAGCCATTCAGACTTCCGATGCTTACCATTTACGAAACCAAGGGGAGTCT 742
Db 212 LysGlyGlyAlaIleAlaThrThrAlaGlyAlaAargIleAlaAsnAsnThrGlyTyrVal 231
QY 743 TCTTTTCTGACAATACCTTCTCGGATTCGAGCTGCAATTTTACAGAAGCTCGGTG 802
Db 232 ArgPheLeuSerAsnIleAlaSerThrSerGlyGlyAlaIleAspAspGlyThrSer 251
QY 803 ACTATTCTTAATAAGTCTAAAGTTCCTTTATTGACAATAAGTCTCACAGGAGCGAGCTCC 862
Db 252 IleLeuSerAsnAsnLysPheLeuTyrPhe-----GluGlyAsnAlaAla 266
QY 863 TCACAACCGGGGATGATGTCAGGAGTGTCTATCTGTGCTTATAAAATAGTACAGATACT 922
Db 267 LysThrThr-----GlyGlyAlaIleCysAsnThrLysAsnSerGlySerPro 282
QY 923 AAGTCTACCTCACTGGAAATCAGATGTTACTTTCAGCAACAATACATACGCAACAGCG 982
Db 283 GluLeuIleSerAsnAsnLysThrLeuIlePheAlaSerAsnValAlaGluThrSer 302
QY 983 GGAGGAGCTATCTATGTGAAAAGCTCGAAGCTTCGGAGGACTTACCTATTCAGT 1042
Db 303 GlyGlyAlaIleHisAlaLysLysLeuAlaLeuSerSerGlyGlyPheThrGluPheLeu 322
QY 1043 AGAAATAGTGTCAATGGAGGTACAGCTCTCTAAAGTGGAGCCATAGCTACGAGATAGT 1102
Db 323 ArgAsnAsnValSerSerAlaThr---ProLysGlyGlyAlaIleSerIleAspAlaSer 341

QY 1103 GGGAAATTGAGTTTATCCCGCGATAGTGGTGACATTTCTTTTAGGGAATACAGTCACT 1162
 DB |||||
 QY 342 GlyLeuSerLeuSerAlaGluThrGlyAsnIleThrPheValArgAsnThrLeuThr 361
 DB |||||
 QY 1163 TCTACT-----ACTCCTGGAGCAATPAGAAAGTAGTAGTACGACTAGGAACGAGTCAAG 1216
 DB |||||
 QY 362 ThrThrGlySerThrAspThrProLysArgAsnAlaIleAsnIleGlySerAsnGlyLys 381
 DB |||||
 QY 1217 ATGACAGCTTGGCTTCTGCTGCTGTAGAGCCATCTACTTCTATCATCCATACTACA 1276
 DB |||||
 QY 382 PheThrGluLeuArgAlaLysAsnHisThrIlePhePheYrAspProIleThr---- 400
 DB |||||
 QY 1277 GCATCATCCACACAGTGTCTTAAAGTCTTAAGAGTCTCCGCGAGTTCGCA 1336
 DB |||||
 QY 401 --SerGluGlyThrSerSerAspValLeuLysIleAsnAsnGlySerAlaGlyAlaLeu 419
 DB |||||
 QY 1337 CTACAATATACAGGAACATCTTCCAGAGGAGAAAGTATACAGAGCAGAGCGCGCA 1396
 DB |||||
 QY 420 AsnProTyrGlnGlyThrIleLeuPheSerGlyGluThrLeuThrAlaAspGluLeuLys 439
 DB |||||
 QY 1397 GATTCTAAATCTTACTTGAAGCTACTACAGCTCTGTAAGTCTTTTACAGAGTACTCTA 1456
 DB |||||
 QY 440 ValAlaAspAsnLeuLysSerSerPheThrGlnProValSerLeuSerGlyGlyLysLeu 459
 DB |||||
 QY 1457 TCTTTAAACATGAGTACTCTCCAGACTCAGCATCTACTCAACAGGAGATTTCTCT 1516
 DB |||||
 QY 460 LeuLeuGlnLysGlyValThrLeuGluSerThrSerPheSerGlnGluAlaGlySerLeu 479
 DB |||||
 QY 1517 CTCGAAATGGAGTAGGAATCTACTAGAA---CCTGCTGATATCTACGACCAATAACAAT 1573
 DB |||||
 QY 480 LeuGlyMetAspSerGlyThrThrLeuSerThrThrAlaGlySerIleThrIleThrAsn 499
 DB |||||
 QY 1574 TTGTCATTATACATCAGTTCTATAGACGGTCAAGAGCAAGCAAAATAGAACCAAGCT 1633
 DB |||||
 QY 500 LeuGlyIleAsnValAspSerLeuGlyLeuLysGlnProValSerLeuThrAlaLysGly 519
 DB |||||
 QY 1634 ACGTCAAAATCTGACTTTATCTGGAACCATCACTTTATTTGGACCGCGACAGCTTT 1693
 DB |||||
 QY 520 AlaSerAsnLysValIleValSerGlyLysLeuAsnLeuIleAspIleGluGlyAsnIle 539
 DB |||||
 QY 1694 TATGAAATCATAGTTTAAAGAACTCTCAGTCCATGACATCTTAGAGTCAAGCTTCT 1753
 DB |||||
 QY 540 TyrGluSerHisMetPheSerHisAspGlnLeuPheSerLeuLeuLysIleThrValAsp 559
 DB |||||
 QY 1754 GGAACGTGTACAAAGC-----ACCGAGTGTACTCCAGATCCCTATAATAGGTGTAG 1801
 DB |||||
 QY 560 AlaAspValAspThrAsnValAspIleSerSerLeuLeuProValProAlaGluAspPro 579
 DB |||||
 QY 1802 AAATTCATTAACGCTATCAGGAACTTTGGGCGCCCAATGTTTGGGGACAGAGGGCTTCT 1861
 DB |||||
 QY 580 AsnSerGluTyrGlyPheGlnGlyGlnTrpAsn---ValAsnTrpThrThrAspThrAla 598
 DB |||||
 QY 1862 ACGACT-----GCAACCTTCAACTGGACTAAACTGGCTATATCTCTAATCCGAG 1912
 DB |||||
 QY 599 ThrAsnThrLysGluAlaThrAlaThrTrpThrLysThrGlyPheValProSerProGlu 618
 DB |||||
 QY 1913 CGTATCGGCTCTTATGCTCCCTAATAGCTTATGGAATGCATTTATAGATATTAGCTCTCT 1972
 DB |||||
 QY 619 ArgLysSerAlaLeuValCysAsnThrLeuTrpGlyValPheThrAspIleArgSerLeu 638
 DB |||||
 QY 1973 CATTATCTATGAGACTGCACAAAGAGGTTCAGGAGACCTGCTTTTCTGCTGCT 2032
 DB |||||
 QY 639 GlnGlnLeuValGluIleGlyAlaThrGlyMetGluHisLysGlnGlyPheTrpValSer 658
 DB |||||
 QY 2033 GGATTTATCTTAACTTCTCCATAAGGATAGTACAAACACACAGCGGGTTCGCAATTG 2092
 DB |||||
 QY 659 SerMetThrAsnPheLeuHisLysThrGlyAspGluAsnArgLysGlyPheArgHisThr 678
 DB |||||
 QY 2093 ACTGGCGGTATGTCATAGGAGAAACCTCATCATCTTCTTTCAGATAGATCTTCTGCT 2152
 DB |||||
 QY 679 SerGlyGlyTyrValIleGlySerAlaHisThrProLysAspAspLeuPheThrPhe 698
 DB |||||
 QY 2153 GCATTTTGTGCTCTTTTGAAGAGATAGAGACTACTTTGTAGCTAAGATCAAGGTACA 2212

DB |||||
 QY 699 AlaPheCysHisLeuPheAlaAsgAspLysAspCysPheIleAlaHisAsnAsnSerArg 718
 DB |||||
 QY 2213 GTCTACGGAGAACTCTATTACACGACGACACAGAAACC-----TATATC 2257
 DB |||||
 QY 719 ThrTyrGlyGlyThrLeuPhePheLysHisSerHisThrLeuGlnProGlnAsnTyrLeu 738
 DB |||||
 QY 2258 TCTCTT---CCTTGCACAACTACGGCTTGTGTCTTCTTCTTCTTCTTACAGAGATTCT 2314
 DB |||||
 QY 739 ArgLeuGlyArgAlaLysPheSerGluSerAlaIleGluLysPheProArgGluLeuPro 758
 DB |||||
 QY 2315 GTTCTCTTTTTCAGAAACCTTACTACACCCATACCGATACGATCTGAAACCAAGAT 2374
 DB |||||
 QY 759 LeuAlaLeuAspValGlnValSerPheSerHisSerAspAsnArgMetGluThrHisTyr 778
 DB |||||
 QY 2375 ACAACATATCTACTGTAAAGGAAGCTGGGAATGATAGTTCTGCTTTAGCAATTCGT 2434
 DB |||||
 QY 779 ThrSerLeuProGluSerGluGlySerTrpSerAsnGluCysIleAlaGlyGlyIleGly 798
 DB |||||
 QY 2435 GGAAGAGCTCGATTTCCTTA---GATGAAAGTGTCTTATTGAGCAGTACATGCCCTTC 2491
 DB |||||
 QY 799 LeuAspLeuProPheValLeuSerAsnProHisProLeuPheLysThrPheIleProGln 818
 DB |||||
 QY 2492 ATGAAATTCAGAGTTTGTCTATGCATCAGCAAGGTTTAAAGACAGGACAGCAAGCT 2551
 DB |||||
 QY 819 MetLysValGluMetValTyrValSerGlnAsnSerPhePheGluSerSerAspGly 838
 DB |||||
 QY 2552 CGTGAATTTGGAAGTACCCCTTGTGAATCTTGCCTTACCTATCCGATCCGATTTGAT 2611
 DB |||||
 QY 839 ArgGlyPheSerIleGlyArgLeuLeuAsnLeuSerIleProValGlyAlaLysPhe--- 857
 DB |||||
 QY 2612 AAGAATTCAGACTCCCAAGATGCA---ACGTACAATCTAACTCTCTGTTATCTGTTGAT 2668
 DB |||||
 QY 858 ValGlnGlyAspIleGlyAspSerTyrThrTyrAspLeuSerGlyPhePheValSerAsp 877
 DB |||||
 QY 2669 CTTGTTCTGATGAACCCGACTGTACGACCACTCGAATACGATGCGGTGATTTTGGAAA 2728
 DB |||||
 QY 878 ValTyrArgAsnAsnProGlnSerThrAlaThrLeuValMetSerProAspSerTrpLys 897
 DB |||||
 QY 2729 ACCTTCGTCGAAATTTGGCAAGACAGCTTTAGTCTCTGTCGACGGACCAATTTTGC 2788
 DB |||||
 QY 898 IleArgGlyGlyAsnLeuSerArgGlnAlaPheLeuLeuArgGlySerAsnAsnTyrVal 917
 DB |||||
 QY 2789 TTTAACTCAAAATTTTGAAGCCTTTAGCCAAATTTCTTTTGAATTCGCTGCGTATCTCGC 2848
 DB |||||
 QY 918 TyrAsnSerAsnCysGluLeuPheGlyHisTyrAlaMetGluLeuArgGlySerSerArg 937
 DB |||||
 QY 2849 AATTACATCTAGACTTAGGACCAATACCAATTC 2884
 DB |||||
 QY 938 AsnTyrAsnValAspValGlyThrLysLeuArgPhe 949

RESULT 13

H86546
 C:Species: Chlamydomonas pneumoniae, Chlamydia pneumoniae
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 15-Jun-2001
 R:Accession: H86546
 R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;
 Nucleic Acids Res. 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
 A:Reference number: A86491; MUID:20330349; PMID:10871362
 A:Accession: H86546
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-928 <STO>
 A:Cross-references: GB:BA000008; NID:98978822; PIDN:BAA98658.1; GSPDB:GN00142
 A:Experimental source: strain J138
 C:Genetics:
 A:Gene: pnp_11
 C:Superfamily: Chlamydomonas pneumoniae polymorphic outer membrane protein G

Alignment Scores:

1.14e-109 Length: 928

Db 285 IleHisAlaLysLysLeuAlaLeuSerSerGlyGlyPheThrGluPheLeuArgAsnAsn 304
 QY 1052 GTCATAGGAGTACAGCTCTCTAAAGTGGAGCCATAGCTATCGAAGATAGTGGGAATTC 1111
 Db 305 ValSerSerAlaThr---ProLysGlyGlyAlaIleSerIleAspAlaSerGlyGluLeu 323
 QY 1112 AGTTATCCCGCAGTACAGTGTGATCTTTTAGGGAATACAGTCATCTTACT--- 1168
 Db 324 SerLeuSerAlaGluThrGlyAsnIlePheValArgAsnThrLeuThrThrGly 343
 QY 1169 ---ACTCTCTGGACCAATAGAGTAGTAGTACGATAGGAGCGAGTCAAGATGACAGCT 1225
 Db 344 SerThrAspThrProLysArgAsnAlaIleAsnIleGlySerAsnGlyLysPheThrGlu 363
 QY 1226 TTGGTCTCTGCTGTAGGACCATCTACTTCTATGATCCATTAAGTACAGATCATCC 1285
 Db 364 LeuArgAlaAlaLysAsnHisThrIlePhePheTyrAspProIleThr-----SerGlu 381
 QY 1286 ACAACAGTTACAGATGCTTAAAGTTAATGAGACTCCGCGAGATTCGCACTACATAT 1345
 Db 382 GlyThrSerSerAspValLeuLysIleAsnAsnGlySerAlaGlyAlaLeuAsnProTyr 401
 QY 1346 ACAGGAACATCATCTTCACAGGAGAAAGTTATCGAGAGACAGGCGCGAGATTCATAA 1405
 Db 402 GlnGlyThrIleLeuPheSerGlyGluThrLeuThrAlaAspGluLeuLysValAlaAsp 421
 QY 1406 AATCTTACTCGAAGCTACTACAGCTGTAACCTCTTCAGGAGTACTCTATCTTTAAAA 1465
 Db 422 AsnLeuLysSerSerPheThrGlnProValSerSerGlyGlyLysLeuLeuGln 441
 QY 1466 CATGAGTACTCTCGAAGTACAGGCAATCTCAACAGGAGATTCCTCGTCAAGATG 1525
 Db 442 LysGlyValThrLeuGluSerThrSerPheSerGlnGluAlaGlySerLeuLeuGlyMet 461
 QY 1526 GACGTAGAACTACTCTAGAA---CCTGCTGATAGTACGCCATTAACAAATTTGGTCATT 1582
 Db 462 AspSerGlyThrLeuSerThrAlaGlySerIleThrIleThrAsnLeuGlyIle 481
 QY 1583 AACATCAGTCTATAGCGGTGCAAGAGGCAAAATATGAACCAAGTACTGCAAAA 1642
 Db 482 AsnValAspSerLeuGlyLysGlnProValSerLeuThrAlaLysGlyAlaSerAsn 501
 QY 1643 AATCTGACTTATCTGGAAACATCATCTTATGACCGCGGCGGCGTCTTATGAAAT 1702
 Db 502 LysValIleValSerGlyLysLeuAsnLeuIleAspIleGluGlyAsnIleTyrGluSer 521
 QY 1703 CATAGTTTAAAGAAATCTCTAGTCTACGACATCTTAGAGCTCAAAAGTCTCTGGAACGTGA 1762
 Db 522 HisMetPheSerHisAspGlnPheSerLeuLeuLysIleThrValAspAlaAspVal 541
 QY 1763 ACAAGC-----ACCGCAGTGAATCCAGATCCTATAATGGGTGAGAAATCCAT 1810
 Db 542 AspThrAsnValAspIleSerSerLeuIleProValProAlaGluAspProAsnSerGlu 561
 QY 1811 TACGCTATCAGGGAACCTGGGCCCAATCTTTGGGGGACAGGGCTCTACGACT--- 1867
 Db 562 TyrGlyPheGlnGlyGlnTrpAsn---ValAsnTrpThrThrAspThrAlaThrAsnThr 580
 QY 1868 -----GCAACCTCAACTGGAATAAACTGGCTATATTCCTAATCCGAGCGGATCCGC 1921
 Db 581 LysGluAlaThrAlaThrTrpThrLysThrGlyPheValProSerProGluArgLysSer 600
 QY 1922 TCTTTAGTCCCTAATAGTATGGATGATATAGATATAGTATAGTCTCTCCATATCTT 1981
 Db 601 AlaLeuValCysAsnThrLeuTrpGlyValPheThrAspIleArgSerLeuGlnGluLeu 620
 QY 1982 ATGAGACTGCAACAGAGGTTGACGAGGACCGCTGCTTTTGGTGTGCTGGATATCT 2041
 Db 621 ValGluIleGlyAlaThrGlyMetGluHisLysGlnGlyPheTrpValSerSerMetThr 640
 QY 2042 AACTTCTTCCATPAGGATAGTACAAAAACAGACCGCGGTTTCGCCAATTTGAGTGGCGGT 2101
 Db 641 AsnPheLeuHisLysThrGlyAspGluAsnArgLysGlyPheArgHisThrSerGlyGly 660

QY 2102 TATGTCATGAGGAACCTACATCTTGTTCAGATAAGATTTCTTAGTGTGCAATTTGT 2161
 Db 661 TyrValIleGlyGlySerAlaHisThrProLysAspAspLeuPheThrPheAlaPheCys 680
 QY 2162 CAGCTCTTTGGAAGAGATGACAGTACTTTGTCTAGCAATCAAGGTACAGTCTACGGA 2221
 Db 681 HisLeuPheAlaArgAspLysAspCysPheIleAlaHisAsnAsnSerArgThrTyrGly 700
 QY 2222 GGAATCTCTATTACACAGCAACAAACC-----TATATCTCTCT--- 2263
 Db 701 GlyThrLeuPhePheLysHisSerHisThrLeuGlnProGlnAsnTyrLeuArgLeuGly 720
 QY 2264 CTTGCCAAATACGCGCTTGTTCGTTCCTTATGTTCTACAGAGATCTCTGTCTCTTT 2323
 Db 721 ArgAlaLysPheSerGluSerAlaIleGluLysPheProArgGluIleProLeuAlaLeu 740
 QY 2324 TCAGGAACCTTAGCTACACCATACCGATACCGATCTGAAACCAAGTATACAACTAT 2383
 Db 741 AspValGlnValSerPheSerHisSerAspAsnArgMetGluThrHisTyrThrSerLeu 760
 QY 2384 CTTACTGTTAAGGAGCTGGGGAATCATAGTTTCCTTTAGAAATTCGTTGGAAGCT 2443
 Db 761 ProGluSerGluGlySerTrpSerAsnGluCysIleAlaGlyGlyIleGlyLeuAspLeu 780
 QY 2444 CCGATTTTCTTAA---GATGAAAGTCTCTATTTGAGCAGTACATGCCCTTCATGAAATG 2500
 Db 781 PropheValLeuSerAsnProHisProLeuPheLysThrPheIleProGlnMetLysVal 800
 QY 2501 CAGTTTGTCTATGCATCAGAGAGGTTTAAAGACAGGAGACAGAGCTCGTGAATTT 2560
 Db 801 GluMetValTyrValSerGlnAsnSerPhePheGluSerSerAspGlyArgGlyPhe 820
 QY 2561 GGAAGTACGCTCTTGTGAATCTTCCCTTACCTATCCGGATCCGATTTGATAGGAATCA 2620
 Db 821 SerIleGlyArgLeuLeuAsnLeuSerIleProValGlyAlaLysPhe---ValGlnGly 839
 QY 2621 GACTGCCAAGATCA---ACGTACATTAATCTCTTGTGTATATCTGTGGATCTTGTTCGT 2677
 Db 840 AspIleGlyAspSerTyrThrTyrAspLeuSerGlyPhePheValSerAspValTyrArg 859
 QY 2678 ACTAACCCCGACTGTACGACCAACTGCGAATTAGCGGTGATTTCTTGAAACCTTCGGT 2737
 Db 860 AsnAsnProGlnSerThrAlaThrLeuValMetSerProAspSerTrpLysIleArgGly 879
 QY 2738 ACGAATTTGGCAAGCAAGCTTTAGTCTCTCGTCAGGGAACCAATTTTGTCTTAATCA 2797
 Db 880 GlyAsnLeuSerArgGlnAlaPheLeuLeuArgGlySerAsnAsnTyrValTyrAsnSer 899
 QY 2798 AATTTTGGAGCTTTAGCCAATTTTCTTTGAATTTGCTGGTGCATCTCCCAATTACAAT 2857
 Db 900 AsnCysGluLeuPheGlyHisTyrAlaMetGluLeuArgGlySerSerArgAsnTyrAsn 919
 QY 2858 GTAGACTTAGGAGCAAAATACCAATTC 2884
 Db 920 ValAspValGlyThrLysLeuArgPhe 928

RESULT 15

E72130
 polypeptide membrane protein G family CP0761 [imported] - Chlamydomonas pneumoniae
 C:Species: Chlamydomonas pneumoniae, Chlamydomonas pneumoniae
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 15-Jun-2001
 C:Accession: E72130; G81541
 R:Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,
 Nature Genet. 21, 385-389, 1999
 A:Title: Comparative genomes of Chlamydomonas pneumoniae and C. trachomatis.
 A:Reference number: A72000; MUID:99260606; PMID:10192388
 A:Accession: E72130
 A:Molecule type: DNA
 A:Residues: 1-841 <ARN>
 A:Cross-references: GB:AE001586; GB:AE001363; NID:94376263; PIDN:AA018172.1; PID:9437
 A:Experimental source: strain CWL029
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke

C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000
 A: Title: Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39.
 A: Reference number: AB1500; MUID:20150255; PMID:10684935

A: Accession: G81541
 A: Molecule type: DNA
 A: Residues: 1-841 <REA>
 A: Cross-references: GB:AE002235; GB:AE002161; NID:g7189672; PIDN:AAF38561.1; PID:g718967
 A: Experimental source: strain AR39, HL cells

C: Genetics:

A: Gene: pmp_2; CP0761

C: Superfamily: Chlamydochila pneumoniae polymorphic outer membrane protein G

Alignment Scores:

Pred. No.: 2,62e-92 Length: 841
 Score: 1590.50 Matches: 375
 Percent Similarity: 52.54% Conservative: 121
 Best Local Similarity: 39.72% Mismatches: 306
 Query Match: 29.22% Indels: 142
 DB: 2 Gaps: 19

US-09-428-122-1 (1-3000) x E72130 (1-841)

QY 134 ACATTTGCTATTTCCCTTGCTATGCTACCGACAGTTTTCGATTCAAGTCG 193
 DB 15 ThrLeuSerMetSerAsnLeuLeuGlyAlaAlaThrThrGluLeuSerAlaSerAsn 34
 QY 194 AGTTTCGATGGG---AATAAAATGCTAATTTTCAGTTCGAGAGTCAGGAAGATGCT 250
 DB 35 SerPheAspGlyThrThrSerThrThrSerPheSerSerLysThrSerSerAlaThrAsp 54
 QY 251 GGAACCTACCTATTTAAGGAAATGTCACCTAGAAAAATTTCTCGGACAGGCACA 310
 DB 55 GlyThrAsnTyrValPheLysAspSerValValIleGluAsnValProLysThrGlyGlu 74
 QY 311 GCAATCAAAAAAGCTGTTTAAACAAC-----ACTAAGGCGGATTGCTTTTCAGAGT 364
 DB 75 ThrGlnSerThrSerCysPheLysAsnAspAlaAlaGlyAspLeuAsnPheLeuGly 94
 QY 365 AAGCGGAACCTCTATTGTTCCAAACGGTGGATGCGAGGACGTAGCAGGGCTGCTGTT 424
 DB 95 GlyGlyPheSerPheThrPheSerAsnIleAspAlaThrThrAlaSerGlyAlaAlaIle 114
 QY 425 AACAGCAGCGTGTGATGATAATCTACACGTTTATAGGTTTCTTCGCTATCTTTATT 484
 DB 115 GlySerGluAlaAlaAsnLysThrValThrLeuSerGlyPheSerAlaLeuSerPheLeu 134
 QY 485 GCGTCTCTGGAAAGTTCGATACTACCGCAAGAGGCGCTTAGCTCTACGGGTAGC 544
 DB 135 LysSerProAlaSerThrValThrAsnGlyLeuGlyAlaIleAsnValLys---GlyAsn 153
 QY 545 TTGAGTTTGACAAAATATGTCAGTTGCTCTTCAGCAAAAATTTTCAACGGTAATGCG 604
 DB 154 LeuSerLeuLeuAspAsnAspLysValLeuIleGlnAspAsnPheSerThrGlyAspGly 173
 QY 605 GGTGCTATCACCGCAAAAATCTTTTCATTAACAGGGACTACAATGTCAGCTCTGTTTCT 664
 DB 174 GlyAlaIleAsnCys----- 178
 QY 665 GAAATACCTCTCAAGAAAGGCGGAGCGCAATTCAGACTTCGATGCCCTTACCATTACT 724
 DB 178 ----- 178
 QY 725 GGAACCAAGGGAAGTCTCTTTTCTGACATATCTTCTCGGATTCGGAGTGCATTT 784
 DB 178 ----- 178
 QY 785 TTTACAGAGCCCTCGGTACTATTCTAATAATGCTAAAGTTTCCTTTATGACATAAG 844
 DB 179 -----AlaGlySerLeuLysIleAlaAsnAsnLysSerLeuSerPheIle----- 193
 QY 845 GTCACAGGAGGAGCTCTCTCAACACCGGGGATATGTCAGGAGGTCTATCTGTGCTTAT 904

DB 193 ----- 193
 QY 905 AAAAAGTACAGATACATAAGGTCAACCTCACTGGAATCAGATGTTACTCTTCAGCAAC 964
 DB 194 -----Gly 194
 QY 965 AATACATCGACACACGCGGAGGAGCTATCTATGTGAAAAGCTCGAAGTGGCTCCGGA 1024
 DB 195 AsnSerSerThrArgGlyGlyAlaIleHisThrLysAsnLeuThrLeuSerGly 214
 QY 1025 GGAATACCTATTCAGTAGAATAAGTGTCAATGAGGTACAGCTCTCT----- 1072
 DB 215 GlyLeuThrLeuPhe-----GlnGlyAsnThrAlaProThrAlaAlaGly 229
 QY 1073 AAAAGTGGAGCCATAGCTATCGAAGATAGTGGGAATTCAGTTTATCCCGCATAGTGT 1132
 DB 230 LysGlyGlyAlaIleAlaIleAlaAspSerGlyThrLeuSerIleSerGlyAspSerGly 249
 QY 1133 GACATTCCTTTTAGGAATACAGTCACTTCTACTCTCTGGAGC---AATAGAGT 1189
 DB 250 AspIleIlePheGluGlyAsnThrIleGlyAlaThr-----GlyThrValSerHisSer 267
 QY 1190 AGTATCGACTAGCAACGAGTCGAAAGATGACAGCTTTCGTTCTCTCTCTGTAGAGCC 1249
 DB 268 AlaIleAspLeuGlyThrSerAlaLysIleThrAlaLeuArgAlaAlaGlnGlyHisThr 287
 QY 1250 ATCTACTTCTATGATCCATPACTACAGGATCATCCACACAGTTCACAGATGCTTTAAA 1309
 DB 288 IleTyrPheTyrAspProIleThrValThrGlySerThrSerValAlaAspAlaLeuAsn 307
 QY 1310 GTTAATGAGACTCGGCGAGATTCGCTACATATACAGGGAACATCATCTTCACAGCA 1369
 DB 308 IleAsnSerProAspThrGlyAspAsnLysGluTyrThrGlyThrIleValPheSerGly 327
 QY 1370 GAAAGTTATCAGACAGAGCGCGCAGATTTCTAAATAATCTTACTTCGAGGTACTACAG 1429
 DB 328 GluLysLeuThrGluAlaGluAlaLysAspGluLysAsnArgThrSerLysLeuLeuGln 347
 QY 1430 CCTGTAACTCTTCAGAGGACTCTCTATCTTTTAAACATGAGGTGACTCTGCAGACTCAG 1489
 DB 348 AsnValAlaPheLysAsnGlyThrValValLeuLysGlyAspValValLeuSerAlaAsn 367
 QY 1490 GCATTTACTCAACAGGCGAGATTCCTGCTCGAAATGAGCTAGAACTACTCTAGAACCT 1549
 DB 368 GlyPheSerGlnAspAlaAsnSerLysLeuIleMetAspLeuGlyThrSerLeu---Val 386
 QY 1550 GCTGATCTAGCACCTATA-----ACAATTTGGTCATTAACATCAGTTCTATAGACGT 1603
 DB 387 AlaAsnThrGluSerIleGluLeuThrAsnLeuGluIleAsnIleAspSerLeuArgAsn 406
 QY 1604 GCAAGGAAGCAAAAATAGAAACCAAGCTACGTCAAAATAATCTGACTTTATCTGGAACC 1663
 DB 407 GlyLysLysIleLysLeuSerAlaAlaThrAlaGlnLysAspIleArgIleAspArgPro 426
 QY 1664 ATCACTTTATGGACCGCGGCGACGCTTTTATGAAAATCATATAGTTTAAAGAAATCCTCAG 1723
 DB 427 ValValLeuAlaIleSerAspGluSerPheTyrGlnAsnGlyPheLeuAsnGluAspHis 446
 QY 1724 TCCTACGAC---ATCTTAGAGCTCAAGGCTTCTGGAACCTGTAACAGCACCGCGAGTACT 1780
 DB 447 SerTyrAspGlyIleLeuGluLeuAspAlaGlyLysAspIleValIleSerAla---Asp 466
 QY 1781 CCAGATCCTTAATGGTGAGAAATTCATTCAGGCTATCAGGGAACCTTGGGGCCCAATTT 1840
 DB 466 erArgSerIleAspAlaValGlnSerPro-TyrGlyTyrGlnGlyLysTrp---ThrIle 484
 QY 1841 GTTTGGGGGACAGGGCTTCTACAGCTGCACCTCACTGGACTAAAACCTGGCTATATT 1900
 DB 485 AsnTrpSerThrAsp---AspLysLysAlaThrValSerTrpAlaLysGlnSerPheAsn 503
 QY 1901 CCTAATCCGAGGCTATCGGCTCTTTAGTCCCTTAATAGCTTATGGAATGCTATTTAGAT 1960
 DB 504 ProThrAlaGluGlnGluAlaProLeuValProAsnLeuLeuTrpGlySerPheIleAsp 523

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 16, 2003, 10:21:33 ; Search time 27.5 Seconds
(without alignments)
10260.372 Million cell updates/sec

Title: US-09-428-122-1
Perfect score: 5409
Sequence: 1 cgccttacctagtaggt.....tggttgtctaaacactttc 3000

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 255726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool.p/US09428122/runat.16122003.102127.10766/app_query.fasta.1.3143
-DB=SwissProt 41 -QPMT=fastan -SUFFIX=rs -MINMATCH=0.1 -IOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09428122 @CNG 1.1.32 @runat.16122003.102127.10766 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2058	38.0	928	PM10_CHLPN	Q9rb55 chlamydia p
2	1982	36.6	928	PM10_CHLPN	Q92398 chlamydia p
3	1964	36.3	936	PM10_CHLPN	Q92898 chlamydia p
4	1936	35.8	930	PM10_CHLPN	Q92393 chlamydia p
5	1855	34.3	928	PM11_CHLPN	O86164 chlamydia p
6	1580.5	29.2	841	PM10_CHLPN	Q923a1 chlamydia p
7	1444	26.7	1276	PM10_CHLPN	Q92899 chlamydia p
8	1442.5	26.7	922	PM10_CHLPN	Q92895 chlamydia p
9	1377.5	25.5	973	PM10_CHLPN	Q92896 chlamydia p
10	1120.5	20.7	1013	PM10_CHLPN	O84879 chlamydia t
11	1051	19.4	987	PM10_CHLPN	Q9pl45 chlamydia m
12	911	16.8	867	PM10_CHLPN	Q9pl41 chlamydia m
13	863	16.0	878	PM10_CHLPN	O84882 chlamydia t
14	685.5	12.7	1609	PM10_CHLPN	Q9z6u5 chlamydia p
15	667.5	12.3	978	PM10_CHLPN	Q92895 chlamydia p
16	666	12.3	947	PM10_CHLPN	Q92813 chlamydia p
17	666	12.3	1723	PM10_CHLPN	Q92812 chlamydia p
18	665.5	12.3	946	PM10_CHLPN	Q92880 chlamydia p

19	650.5	12.0	952	1	PM16_CHLPN	O9z882 chlamydia p
20	605.5	11.2	975	1	PM16_CHLPN	O84417 chlamydia t
21	604.5	11.2	938	1	PM15_CHLPN	Q92883 chlamydia p
22	597	11.0	1016	1	PM16_CHLPN	O84880 chlamydia t
23	594	11.0	980	1	PM16_CHLPN	Q9pl44 chlamydia m
24	590.5	10.9	976	1	PM16_CHLPN	Q9pl43 chlamydia m
25	576.5	10.7	1531	1	PM16_CHLPN	O84818 chlamydia t
26	559	10.3	1754	1	PM16_CHLPN	O84418 chlamydia t
27	549.5	10.2	976	1	PM16_CHLPN	O9pl47 chlamydia m
28	548.5	10.1	1672	1	PM16_CHLPN	Q9pl42 chlamydia m
29	544	10.1	1520	1	PM16_CHLPN	Q9pl46 chlamydia m
30	541.5	10.0	984	1	PM16_CHLPN	O84877 chlamydia t
31	533.5	9.9	1460	1	PM16_CHLPN	O9pl41 chlamydia m
32	523	9.7	1770	1	PM16_CHLPN	O84419 chlamydia t
33	490	9.1	1034	1	PM16_CHLPN	P38008 chlamydia p
34	475.5	8.8	514	1	PM16_CHLPN	Q923d6 chlamydia t
35	454	8.4	1025	1	PM16_CHLPN	Q9pl46 chlamydia m
36	269	5.0	2249	1	PM16_CHLPN	P15921 rickettsia
37	251	4.6	2021	1	PM16_CHLPN	Q52657 rickettsia
38	238.5	4.4	1025	1	PM16_CHLPN	P35828 caulobacter
39	237.5	4.4	881	1	PM16_CHLPN	P47033 saccharomyc
40	234	4.3	1645	1	PM16_CHLPN	P96989 r outer mem
41	233.5	4.3	1609	1	PM16_CHLPN	P25853 saccharomyc
42	232	4.3	1643	1	PM16_CHLPN	Q53020 r outer mem
43	230	4.3	995	1	PM16_CHLPN	P40442 saccharomyc
44	230	4.3	1250	1	PM16_CHLPN	P45508 escherichia
45	228	4.2	3178	1	PM16_CHLPN	Q09624 caenorhabdi

ALIGNMENTS

RESULT 1
PM10_CHLPN
ID PM10_CHLPN STANDARD; PRT; 928 AA.
AC Q9RB65; O86163; Q9RB64; Q9S6P2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable outer membrane protein pmpl0 precursor (Polymorphic membrane protein 10) (Outer membrane protein 5).
GN PM10 OR CMPS OR CP0303.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VR1310;
RX MEDLINE=20007584; PubMed=10539856;
RA Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT "Molecular biology of Chlamydia pneumoniae surface proteins and their role in immunopathogenicity.";
RL Am. Heart J. 138:S491-S495(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoko M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CML029 from USA.";


```

RL Nucleic Acids Res. 28:2311-2314(2000).
RN [4]
RP SEQUENCE OF 1-914 FROM N.A.
RC STRAIN=CWL029/VR-1310;
RX MEDLINE=99081766; PubMed=9864239;
RA Knudsen K., Madsen A.S., Mygind P., Christiansen G., Birkelund S.;
RT "Identification of two novel genes encoding 97- to 99-kilodalton outer
RT membrane proteins of Chlamydia pneumoniae.";
RT Infect. Immun. 67:375-383(1999).
CC -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC (POTENTIAL)
CC -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: AJ133034; CAB37071.1; -
CC EMBL: AF002192; AAF38160.1; -
CC EMBL: AF002546; BAA98657.1; -
CC EMBL: AJ001311; CAA04671.1; -
CC PIR: G81591; G81591.
CC PIR: G86546; G86546.
CC PHCI-2DPAGE; O86163; -.
CC TIGR: CP0303; -.
CC InterPro: IPR006315; Autotransport.
CC InterPro: IPR003368; Chlamydia_PMP.
CC Pfam: PF02415; DUF145; 2.
CC TIGRFAMs: TIGR01414; autotrans_bar1; 1.
CC TIGRFAMs: TIGR01376; POMP repeat; 6.
CC Outer membrane; Signal; Multigene family.
CC SIGNAL 1 25 POTENTIAL.
CC CHAIN 26 928 PROBABLE OUTER MEMBRANE PROTEIN PMP10.
CC SEQUENCE 928 AA; 97229 MW; 0590D5206A1DD0E1 CRC64;
CC
CC Alignment Scores:
CC Pred. No.: 2,22e-116 Length: 928
CC Score: 2058.00 Matches: 444
CC Percent Similarity: 61.68% Conservative: 142
CC Best Local Similarity: 46.74% Mismatches: 320
CC Query Match: 38.05% Indels: 44
CC DB: 1 Gaps: 20
CC
CC US-09-428-122-1 (1-3000) x PMP10_CHLPP (1-928)
CC
CC 101 ATGAAGTCTCTTCCCAAGTTTGATTTT---TCTACATTTGCTATTTTC----- 148
CC 1 MetLysSerGlnPheSerTrpLeuValLeuSerSerThrLeuAlaCysPheThrSerCys 20
CC
CC 149 CCTTTCTCTATGATCTCTACCGACAGACAGTTTTCGATTCAAGTCGCGAGTTTCGATGGGAAT 208
CC 21 SerThrValPheAlaAlaThrAlaGluAsnIleGlyProSerAspSerPheAspGlySer 40
CC
CC 209 AAAAAT---GGTAATTTTTCAGTCTGAGAGTCGAGAGTCAAGATGCTCGAACTACCTACCTTA 265
CC 41 ThrAsnThrGlyThrThrThrProLysAsnThr-----ThrThrGlyIleAspTyrThr 58
CC
CC 266 TTTAAGGGAATGCTACTAGAAAATATTCCTGGAAACAGGCACAGCATCACAAAAAGC 325
CC 59 LeuThrGlyAspIleThrGluGlnAsnLeu---GlyAspSerAlaAlaLeuThrLysGly 77
CC
CC 326 TGTTTTAAACAACAAAGGCGGATTTGACTTTTCACAGGTAACGGGAACTCTATTGTTTC 385
CC 78 CysPheSerAspThrThrGluSerLeuSerPheAlaGlyLysGlyTyrSerLeuSerPhe 97
CC
CC 386 CAAACCGTGATGTCAGGAGCTGTAGCAGGGCTGCTGTTAAACAGCAGCGCTGGTAGATAAA 445
CC 98 LeuAsnIleLysSer---SerAlaGluGlyAlaAlaLeu---SerValThrThrAspLys 115

```

Db	464	ThrAlaGlySerSerValIleMetAspAlaGlyThrThrLeuLysAlaSerThrGluGlu	483
Qy	1559	AGCACCAATAACAATTTGGTCATTAAACATCAGTTCTATAGACGGTCAAAAGCGCAAA	1618
Db	484	ValThrLeuThrGlyLeuSerIleProValAspSerLeuGlyGluGlyLysValVal	503
Qy	1619	ATGAACACCAAGCTACGTCACAAAATCTCGATTTATCTGGAACCATCACCCTATTATCGAC	1678
Db	504	IleAlaAlaSerAlaAlaSerLysAsnValAlaLeuSerGlyProIleLeuLeuLeuAsp	523
Qy	1679	CCGACGGGCAGCTTTTATGAAAAATCATAGTTTAAAGAAATCCTCAGTCCTACGACATCTTA	1738
Db	524	AsnGlnGlyAsnAlaTyrgluAsHisAspLeuGlyLysThrGlnAspPheSerPheVal	543
Qy	1739	GAGCTCAAAAGCTTCGGAACCTGTAAACAAGCACCCAGTCAGCTCCAGATCCTATAATGGGT	1798
Db	544	GlnLeuSerAlaLeuGlyThrAlaThrThrThrAspValProAlaValProThrValAla	563
Qy	1799	GAGAAATTCATATACGGCTATCAGGAACTTTGGGGCCCAATTTGTTGG-----	1846
Db	564	ThrProThrHisTyrglyTyrglnGlyThrTripoly--MetThrTripValAspAspThr	582
Qy	1847	---GGGACAGGGGCTTCTACGACTCGAACCTTCACTGGACTAAACATGGCTATATTCCT	1903
Db	583	AlaSerThrProLysThrLysThrAlaThrLeuAlaTriphThrAenthrglyTyzLeuPro	602
Qy	1904	AATCCCAGAGCGTATCGGCTCTTATAGTCCCTAAATAGCTTATGGATGCATTTATAGATATT	1963
Db	603	AsnProgluargGlnGlyProLeuValProAsnSerLeuTripglySerPheSerAspIle	622
Qy	1964	AGTCTCTCCATATCTTATGGAGCTGCACAAACGAAAGGTTGCGAGGAGACCGTGCTTTT	2023
Db	623	GlnAlaIleGlnGlyValIleGluargSerAlaLeuThrLeuCysSerAspArgGlyPhe	642
Qy	2024	TGGTGCTCGATATATCACTTCTTCCATAAGATAGTACAAAAACACACGCGGGTTT	2083
Db	643	TriAlaAlaGlyValAlaAsnPheLeuAspLysAspLysGlyGluLysArgLysTyz	662
Qy	2084	CGCCATTGAGTGGCGGTATGTCTATAGGAGGAAACCTACATATCTGTCAGATAAGATT	2143
Db	663	ArgHisLysSerGlyGlyTyzAlaIleGlyAlaAlaGlnThrCysSerGluAsnLeu	682
Qy	2144	CTTAGTGCTCATTTGCTCAGCTCTTTGGAGAGATAGAGACTACTTTGTAGCTAAGAAT	2203
Db	683	IleSerPheAlaPheCysGlnLeuPheGlySerAspLysAspPheLeuValAlaLysAsn	702
Qy	2204	CAAGGTACGCTACGGAGGAACCTCTATATACCAGCACAAACGAAACCTATATCTCTCT	2263
Db	703	HisThrAspThrTyzAlaGlyAlaPheTyzIleGlnHis-----IleThrGlu	718
Qy	2264	CCTTGCAAACTACGGCCTTGTCGTTCTTATCTTCTCT-----ACAGAGATTCCCT	2314
Db	719	CysSerGlyPheIleGlyCysLeuLeuAspLysLeuProglySerTripSerHisLysPro	738
Qy	2315	GTTCCTCTTTTACGAAACCTTAGCTACACCCATACCGGATAACGATCTGAAAAACCAAGTAT	2374
Db	739	LeuValLeuGluGlyGlnLeuAlaTyzSerHisValSerAsnAspLeuLysThrLysTyz	758
Qy	2375	ACAACATATCTACTGTTAAAGAGCTGGGGATCATAGTTTCGCTTTAGAAATTCGGT	2434
Db	759	ThrAlaTyzProGluValLysGlySerTripglyAsnAsnAlaPheAsnMetMetLeuGly	778
Qy	2435	GGAAGAGTCGATTTGCTTAGATGAAGTGCTCTATTTCAGCAGTACATGCCCTTCATG	2494
Db	779	AlaSerSerHisSerTyzProGluTyzLeuHisCysPheAspThrTyzAlaProTyzIle	798
Qy	2495	AAATTGCGATTTGTCTATGCAATCAGGAAGAGTTTAAAGAACAGGGAACAGAAAGCTCGT	2554
Db	799	LysLeuAsnLeuThrTyzIleArgGlnAspSerPheSerGluLysGlyThrGluGlyArg	818
Qy	2555	GAATTTGGAGTAGTCGCTTTGTGGAATCTTTCCTTACCTATCGGATTCGATTTGATAAG	2614

819	SerPheAspSerAsnLeuPheAsnLeuSerLeuProIleGlyValIysPheGluLys	838
2615	GAATCAGACTGCAGAGTCAAGCTCAACTCAACTCTCGTTATATACGTGGATCTTGTT	2674
839	PhSerAspCysAsnAspPheSerTyrAspLeuThrLeuSerTyrValProAspLeuIle	858
2675	CGTAGTAACCCCGACTGTACGACAACACGCTAGCGGTGATCTTTCGAAACCTTC	2734
859	ArgAsnAspProLysCysThrThrAlaLeuValIleSerGlyAlaSerTrpGluThrTyr	878
2735	GGTACGAATTGCGACAGACAGCTTGTAGTCTCTTCGTCAGGAGCAACATTTTTCGTTAAC	2794
879	AlaAsnAsnLeuAlaArgGlnAlaLeuGlnValArgAlaGlySerHisTyrAlaPheSer	898
2795	TCAAAATTTTCAAGCCTTTAGCCAAATTTCTTTGAAATTTGGTGGGTGCATCTCGCAATTAC	2854
899	ProMetPheGluValLeuGlyGlnPheValPheGluValArgGlySerSerArgIleTyr	918
2855	AATGTAGACTTAGAGCAGCAAAATCAATTC	2884
919	AsnValAspLeuGlyGlyLysPheGlnPhe	928
RESULT 2		
PMP3_CHLPN		
ID	PMP3_CHLPN	STANDARD; PRT; 928 AA.
AC	Q9Z398;	
DT	16-OCT-2001 (Rel. 40, Created)	
DT	16-OCT-2001 (Rel. 40, Last sequence update)	
DT	16-OCT-2001 (Rel. 40, Last annotation update)	
DE	Probable outer membrane protein pmp9 precursor (Polymorphic membrane protein 9) (outer membrane protein 10).	
DE	GNP OR OMP10 OR CPN0447 OR CP0306.	
OS	Chlamydia pneumoniae (Chlamydophila pneumoniae).	
OC	Bacteria; Chlamydiae; Chlamydophila pneumoniae.	
OX	NCBI_TaxID=83558;	
UN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=VR1310;	
RC	MEDLINE=20007594; PubMed=10539856;	
RA	Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,	
RA	Madsen A.S., Knudsen K., Falk E., Birkelund S.;	
RA	"Molecular biology of Chlamydia pneumoniae surface proteins and their role in immunopathogenicity";	
RT	Am. Heart J. 138:S491-S495(1999).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=CWL029;	
RC	MEDLINE=99206606; PubMed=10192388;	
RA	Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,	
RA	Olinger L., Grimwood J., Davis R.W., Stephens R.S.;	
RT	"Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";	
RN	Nat. Genet. 21:385-389(1999).	
UN	[3]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=AR39;	
RC	MEDLINE=20150255; PubMed=10684935;	
RA	Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,	
RA	White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,	
RA	Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,	
RA	Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,	
RA	Eisen J., Fraser C.M.;	
RT	"Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.";	
RL	Nucleic Acids Res. 28:1397-1406(2000).	
UN	[4]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=J138;	
RC	MEDLINE=20330349; PubMed=10871362;	
RA	Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,	
RA	Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;	
RT	"Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CWL029 from USA.";	
RL	Nucleic Acids Res. 28:2311-2314(2000).	


```

Db      506 AlaThrGlnAlaSerGlnThrValThrLeuSerGlySerLeuValAspProSer 525
QY      1685 GGCACGTTTATGAAATCATAGTTTAAAGAAATCCTCAGTCCTCAGACATCTTAGAGCTC 1744
Db      526 GlyAsnValTyrGluAspValSerTrpAsnAsnProGlnValPheSerLeuThrLeu 545
QY      1745 AAAGCTTCT-----GGAAGTGTAAACAGCAGCGAGTCACTCCAGATCCTATAATGGGT 1798
Db      546 ThrAlaAspProAlaAsnIleHisIleThrAspLeuAlaAlaAspProLeuGluLys 565
QY      1799 GAGAAATTCATTACCGCTTATCAGGGAACCTTGGGGCCCAATTTGGGGGAGCAGGGGCT 1858
Db      566 AsnProIleHisTrpGlyTyrGlnGlyAsnTrpAla---LeuSerTrpGlnGluAspThr 584
QY      1859 TCTACG-----ACTGCAACCTTCAACTGAGCTAAACAGCGCTATATCTCTAATCC 1909
Db      585 AlaThrLysSerLysAlaAlaThrLysValArgGlnSerGlnGluThrArgGlyIleTrpCys 604
QY      1910 GAGCGTATCGGCTCTTTAGTCCCTAATAGCTTATGGAATGCATTTATAGATATTAGCTCT 1969
Db      605 GluArgArgGlyThrLeuValAlaAsnThrLeuTrpGlySerPheValAspValArgSer 624
QY      1970 CTCATTATCTTATGAGAGCTGCAACAGGAGGCTTGCAGGAGACCGTCTTTTGGTGT 2029
Db      625 IleGlnGlnLeuValAlaThrLysValArgGlnSerGlnGluThrArgGlyIleTrpCys 644
QY      2030 GCTGATATCTAATCTTCCATAGGATAGTACAAACACGAGCGGGTTCGCCAT 2089
Db      645 GluGlyIleSerAsnPhePheHisLysAspSerThrLysIleAsnLysGlyPheArgHis 664
QY      2090 TTGAGTGGCGGTATGCTATAGAGGAAACCTACATCTTGTTCAGATAAGATCTTAGT 2149
Db      665 IleSerAlaGlyTyrValGlyAlaThrThrLeuAlaSerAspAsnLeuIleThr 684
QY      2150 GTCGATTTTGTAGCTTTTGGAGAGATAGAGACTACTTTGTAGTAAAGATCAAGGT 2209
Db      685 AlaAlaPheCysGlnLeuPheGlyLysAspArgAspHisPheIleAsnLysAsnArgAla 704
QY      2210 ACAGTCTACGGAGGAACTCTTATACAGCAGCAGCAAGAACTATATCTCTCTCTGTC 2269
Db      705 SerAlaTyrAlaAlaSerLeuHisLeuGlnHisLeuAlaThrLeuSerSer----- 721
QY      2270 AAATACGGCTTGTGCTGTCTATGTTCTCT-----ACAGAGATTCCTGTTCTC 2320
Db      722 -----ProSerLeuLeuArgTyrLeuProGlySerGlySerGlnProValLeu 738
QY      2321 TTTTCAGGAAACCTTAGCTACACCATACGATACGATCTGAAACCAAGTATACACA 2380
Db      739 PheAspAlaGlnIleSerTyrIleTyrSerLysAsnThrMetLysTyrTyrThrGln 758
QY      2381 TATCTCTACTTTAAAGGAGCTGGGGAGTATGATGTTGCTTTTGAATTCGGTGGAGA 2440
Db      759 AlaProLysGlyGlySerTrpTyrAsnAspGlyCysAlaLeuGluLeuAlaSerSer 778
QY      2441 GCTCGG---ATTGTCTAGATCAAAAGTGTCTATTGAGCAGTACATGCTCTCATGAA 2497
Db      779 LeuProHisThrAlaLeuSerHisGlyGlyLeuPheHisAlaTyrPheProPheIleLys 798
QY      2498 TTGCAGTTTGTTCATGCATCAGGAAGGTTTAAAGAACAGCAAGCAAGAA---GCTCGT 2554
Db      799 ValGluAlaSerTyrIleHisGlnAspSerPheLysGluArgAsnThrThrLeuValArg 818
QY      2555 GNAATTGGAAGTAGCGGCTTCTGAAATCTTGCCTTACCTATCGGATCGGATTCGATAAG 2614
Db      819 SerPheAspSerGlyAspLeuIleAsnValSerValProIleGlyIleThrPheGluArg 838
QY      2615 GAATCAGACTGCAAGATGCAAGTACATCTAATCTTGTGTTACTGTGATCTTGT 2674
Db      839 PheSerArgAsnGluArgAlaSerTyrGluAlaThrValIleTyrValAlaAspValTyr 858
QY      2675 CGTAGTACCCGAGCTGTACGCAACACTGCGAATTAGCGTGTGATCTTGGAAAAACCTTC 2734

```

```

Db      859 ArgLysAsnProAspCysThrThrAlaLeuLeuIleAsnAsnThrSerTrpLysThrThr 878
QY      2735 GGTACGAATTTGCGAAGACAGACTTTAGTCTCTTCGTGCGAGGAACCATTTTGTCTTAA 2794
Db      879 GlyThrAsnLeuSerArgGlnAlaGlyIleGlyArgAlaGlyIlePheTyrAlaPheSer 898
QY      2795 TCAAAATTTGAGACCTTTAGCCAAATTTCTTTTGAATTCGCTGGTCAATCTCGCAATTAC 2854
Db      899 ProAsnLeuGluValThrSerAsnLeuSerMetGluIleArgGlySerSerArgSerTyr 918
QY      2855 AATGTAGACTTAGGACCAAAATACCAATTC 2884
Db      919 AsnAlaAspLeuGlyGlyLysPheGlnPhe 928

RESULT 3
PMP7_CHLPN
ID PMP7_CHLPN STANDARD; PRT; 936 AA.
AC Q9Z898; Q9Z842; Q9Z504;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Probable outer membrane protein pmp7 precursor (Polymorphic membrane protein 7) (Outer membrane protein 12).
DE PMP7 OR OMP12 OR CPN0445 OR CP0308.
OS Chlamydia pneumoniae (Chlamydiophila pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydiophila.
OX NCBI_TaxID=83358;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative Genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RC SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RC SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi P., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [4]
RC SEQUENCE OF 658-936 FROM N.A.
RC STRAIN=VR1310;
RX MEDLINE=20007584; PubMed=10539856;
RA Christiansen G., Bosen T., Hjerno K., Daugaard L., Mygind P.,
RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT "Molecular biology of Chlamydia pneumoniae surface proteins and their role in immunopathogenicity.";
RL Am. Heart J. 138:S491-S495(1999).
CC -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES) (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>) or send an email to license@isb-sib.ch).

```

EMBL; AE001627; RAD18589.1; -.
EMBL; AE002193; AAF38165.1; -.
EMBL; AF0002546; BAA98653.1; -.
EMBL; AJ133034; CAB37067.1; -.
PIR; B81591; B81591.
PIR; C72078; C72078.
PIR; C86545; C86546.
PDB1-2DPAGE; Q9Z898; -.
TIGR; CP0308; -.
InterPro; IPR006315; Autotransport.
InterPro; IPR005546; Chlamiporter.
InterPro; IPR003368; Chlamydia_PMP.
Pfam; PF03797; Autotransporter_1.
Pfam; PF02415; DUF145; 2.
TIGRfam; TIGR01414; autotrans_bar1; 1.
TIGRfam; TIGR01376; POMP repeat; 6.
Outer membrane; Signal; Multigene family; Complete proteome.
SIGNAL 1 23
CHAIN 24 936
CONFLICT 658
CONFLICT 822
SEQUENCE 936 AA; 100105 MW; 33881DB3C950AF95A CRC64;
PROBABLE OUTER MEMBRANE PROTEIN PMP7.
PRHGFGRH1 -> EDNIRYRN (IN REF. 4).
Y -> H (IN REF. 1 AND 4).

```

```
Db 531 TyrlGluAsnHisAsnLeuLysSerAlaSerThrTyrlProLeuLeuGluLeuThrAla 550
QY 1748 ---GCTCTGGAACGTAAACAGCACCAGCAGTCCAGATCCCTATATATGGTGGTACAGAA 1804
Db 551 GlyAlaAsnGlyThrLeuLeuGluLeuGluLeuGluLeuGluLeuGluLeuGluLeu 570
QY 1805 TTCCATTAGCGTATCAGGAGCACTTGGGGCCCAATTTGGGGACAGGGGCTTCTACG 1864
Db 571 ThrHisTyrlGlyTyrlGlnGlyAsnTrp---GlnLeuSerTrpAlaAsnAlaThrSerSer 589
QY 1865 ---ACTGCAACCTCAACTGGAGTAAACTGGCTATATTCCTTAATCCCGAGGTATCGGC 1921
Db 590 LysIleGlySerIleAsnTrpThrArgThrGlyTyrlProSerProGluArgLysSer 609
QY 1922 TCTTTAGTCCCTAATAGCTTATGGAATGCAATTTATAGATATAGTCTCTCCATATCTT 1981
Db 610 AsnLeuProLeuAsnSerLeuTrpGlyAsnPheIleAspIleArgSerIleAsnGlnLeu 629
QY 1982 ATGGAGACTGCAAAACGAAGGGTTGCAGGAGACCGTCTTTTGGTGTGGATTATCT 2041
Db 630 IleGluThrLysSerSerGlyGluProPheGluArgGluLeuTrpLeuSerGlyIleAla 649
QY 2042 AACTTCTTCCATAGGATAGTACAAAACACAGACCGGGTTTCGCCATTCAGTGGCGGT 2101
Db 650 AsnPhePheTyrlArgAspSerMetProThrArgHisGlyPheArgHisIleSerGlyGly 669
QY 2102 TATGTCATAGGAGGAAACCTACATCTTGTTCAGATAGATCTTAGTGTGCTGCTATTTGT 2161
Db 670 TyrlAlaLeuGlyIleThrAlaThrProAlaGluAspGlnLeuThrPheAlaPheCys 689
QY 2162 CAGCTCTTTGGAAGAGATAGAGACTACTTGTAGCTAAGATCAAGGTACAGTCTACGGA 2221
Db 690 GlnLeuPheAlaAargAspAsnHisIleThrGlyLysAsnHisGlyAspThrTyrlGly 709
QY 2222 GGAATCTCTATTACACCAACAAACCTATATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2275
Db 710 AlaSerLeuTyrlPheHisHisThrGluGlyLeuPheAspIleAlaAsnPheLeuTrpGly 729
QY 2276 -----CGGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2323
Db 730 LysAlaThrArgAlaProTrpValLeuSerGluLeuSerGlnIleIleProLeuSerPhe 749
QY 2324 TCAGGAACCTTAGCTACACCATACGATACAGTCTGAAACCAAGTATACACATAT 2383
Db 750 AspalalysPheSerTyrlLeuHisThrAspAsnHisMetLysTyrlTyrlThrAspAsn 769
QY 2384 CCTACTGTTAAGGAAGCTGGGGAATGATAGTTTCGCTTTAGAAATTCGGTGGAGAGCT 2443
Db 770 SerIleIleLysGlySerTrpArgAsnAspAlaPheCysAlaAspLeuGlyAlaSerLeu 789
QY 2444 CCGATTTGCTTAGATGAAGTCTCTATTGAGCAGTACATGCCCTTCATGAATTCGAG 2503
Db 790 ProPheValIleSerValProTyrlLeuLysGluValGluProPheValLysValGln 809
QY 2504 TTTGCTATGACATCAGGAAGGTTTAAAGAACAGAGGACAGAGGTCGTGAATTTGGA 2563
Db 810 TyrlleTyrlAlaHisGlnAspPheTyrlGluArgTyrlAlaGluGlyArgAlaPheAsn 829
QY 2564 AGTAGCGCTGTTGGAATCTTGCCTTACCTATCCGATCCGATTCGATTAAGGAATCAGAC 2623
Db 830 LysSerGluLeuIleAsnValGluIleProIleGlyValThrPheGluArgAspSerLys 849
QY 2624 TGCCAAGATCAAGTCAACTACTCTGTTGTTATCTGTTGATCTTGTTCGTAGTAAC 2683
Db 850 SerGluLysGlyThrTyrlAspLeuThrLeuMetTyrlleLeuAspAlaTyrlArgAsn 869
QY 2684 CCCGACTGTACGACAACTACGATAGCGGTGATCTTGGAAACCTTCGGTACCAAT 2743
Db 870 ProlysCysGlnThrSerLeuIleAlaSerAspAlaAsnTrpMetAlaTyrlGlyThrAsn 889
QY 2744 TTGCAAGACAGACTTTAGTCTCTGTCGAGGGAACCATTTTCTGCTTAACCTCAAAATTT 2803
```

```
Db 890 LeuAlaArgGlnGlyPheSerValArgAlaAlaAsnHisPheGlnValAsnProHisMet 909
QY 2804 GAAGCCTTTAGCCAAATTTCTTTGAATGGTGGTGCATCTCGCAATCAATGATGAC 2863
Db 910 GluTrpPheGlyGlnPheAlaPheGluValArgSerSerArgAsnTyrlAsnThrAsn 929
QY 2864 TTAGAGCAAAATACCAATTC 2884
Db 930 LeuGlySerLysPheCysPhe 936
RESULT 4
PMP8_CHLPN
ID_PMP8_CHLPN STANDARD; PRT; 930 AA.
AC Q92393; Q9RB66;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable outer membrane protein pmp8 precursor (Polymorphic membrane
protein 8) (Outer membrane protein 11).
GN PMP8 OR OMP11 OR CPN0446 OR CP0307.
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VR1310;
RX MEDLINE=20007584; PubMed=10539856;
RA Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT "Molecular biology of Chlamydia pneumoniae surface proteins and their
role in immunopathogenicity".
RL Am. Heart J. 138:S491-S495(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=39206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.".
RL Nat. Genet. 21:385-389(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Winn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
pneumoniae AR39.".
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
from Japan and CWL029 from USA.".
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -1- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
(POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

```

DR EMBL; AJ133034; CAB37068.1; -.
DR EMBL; AB001627; AAD18590.1; -.
DR EMBL; AB002193; AAF38164.1; -.
DR EMBL; AF002546; BAA98654.1; -.
DR PIR; A81591; A81591.
DR PIR; D72078; D72078.
DR PHCI-2DPAGE; Q92393; -.
DR TIGR; CP0307; -.
DR InterPro; IPR006315; Autotransport.
DR InterPro; IPR005546; Autotransporter.
DR InterPro; IPR003368; Chlamydia PMP.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF02415; DUF145; 2.
DR TIGRfams; TIGR01414; autotrans_bar1; 1.
DR TIGRfams; TIGR01376; POMP repeat; 6.
DR Outer membrane; Signal; Multigene family; Complete proteome.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 930 PROBABLE OUTER MEMBRANE PROTEIN PMP8.
FT CONFLICT 177 177 T -> A (IN REF. 3 AND 4).
SQ SEQUENCE 930 AA; 97669 MW; 46A9B5EBB913C4C CRC64;

Alignment Scores:
Pred. No.: 4,75e-109 Length: 930
Score: 1936.00 Matches: 415
Percent Similarity: 60.86% Conservative: 165
Best Local Similarity: 43.55% Mismatches: 325
Query Match: 48 Indels: 48
DB: 1 Gaps: 17

US-09-428-122-1 (1-3000) x PMP8_CHLPN (1-930)
QY 101 ATGAAGTCCTCTTCCCGAGTTGGTATTTCTACATTGCTATTTTCCTTTGTCTATG 160
Db 1 MetLysileProLeuHisLysLeuLeuLeuSerSerThrLeuValThrProIleLeuLeu 20
QY 161 ---ATTGCTACC-----GAGACAGTTTTCGATTTCAGTCGAGTTTCGATGGGAAT 208
Db 21 SerIleAlaThrTyrglyAlaAspAlaSerLeuSerProThrAspSerPheAspGlyAla 40
QY 209 AAAAATGGTAATTTTCAGTTCGTGAGTCAGACAGAGATGCTGGAAGTACCTACCTATTT 268
Db 41 GlyGlySerThrPheThrProLysSerThrAlaAspAlaAsnGlyThrAsnTyrrValLeu 60
QY 269 AAGGGAATCTACTAGAAATATCTCGGACAGGACAGACATCAAAAGCTGT 328
Db 61 SerGlyAsnValTyrrile---AsnAspAlaGlyLysGlyThrAlaLeuThrGlyCysCys 79
QY 329 TTTTAAACAACACTAAGGGCGATTGACTTTTCACAGGTAAACGGGAACCTCTATTGTTCCAA 388
Db 80 PheThrGluThrThrGlyAspLeuThrPheThrGlyLysGlyTyrrSerPheSerPheAsn 99
QY 389 ACGGTGAGTCAGGAGCTGTAGCAGGGCTGCTGTTAAACAGACGGTGGTAGATAATCT 448
Db 100 ThrValAspAlaGlySerAsnAlaGlyAlaAlaAla---SerThrThrAlaAspLysAla 118
QY 449 ACCACGTTTATAGGTTTCTTCGCTATCTTTTATTGCGTCTCTCGAAGTTCGATAACT 508
Db 119 LeuThrPheThrGlyPheSerAsnLeuSerPheIleAlaAlaProGlyThrThrValAla 138
QY 509 ACGGCAAAAGGACCGTGTAGCTGCTCTAGCGGTAGCTTGAGTTTGCACAAAATGTCAGT 568
Db 139 SerGlyLysSerThrLeuSer---SerAlaGlyAlaLeuAsnLeuThrAspAsnGlyThr 157
QY 569 TTGCTCTTCAGCAAAACTTTTCAACGGAT-----AATGGCGGTCTATCACCGCA 619
Db 158 IleLeuPheSerGlnAsnValSerAsnGlnAlaAsnAsnAsnGlyGlyAlaIleThrThr 177
QY 620 AAAACTCTTTTCAATTAACAGGACTACAATGTACGCTCTGTTTCTGAAATACCTCTCA 679
Db 178 LysThrLeuSerIleSerGlyAsnThrSerSerIleThrPheThrSerAsnSerAlaLys 197
QY 680 AAGAAAGGGGAGCCCATTCAGACTTCGATGCCCTTACCATTACTGCAAAACCAAGGGGAA 739
Db 197 LysThrLeuSerIleSerGlyAsnThrSerSerIleThrPheThrSerAsnSerAlaLys 197

```

```

Db 198 LysLeuGlyGlyAlaIleTyrrSerSerAlaAlaAlaSerIleSerGlyAsnThrGlyGln 217
QY 740 GTCTCTTTCTGACAATACTTCTTCGGATTCTGGAGCTGCAATTTTTCAGAAAGCTCG 799
Db 218 LeuValPheMetAsnAsnLysGlyGluThrGlyGlyGlyAlaLeuGlyPheGluAlaSer 237
QY 800 GTGACTATTTCTAAATCTAAAGTTTCTTTTATGCAATAAGGTACACGAGCAGAC 859
Db 238 SerSerIleThrGlnAsnSerLeuPhePheSerGlyAsnThrAlaThrAspAlaAla 257
QY 860 TCTCTCAACAACGGGGATATGTCAGAGGTGCTATCTGCTGCTTATAAACTAGTACAGT 919
Db 258 Gly-----LysGlyGlyAlaIleTyrrCysGluLysThrGlyGluThr 271
QY 920 ACTAAGTCCACCTCACTGGAATCAGATGTTACTTCTCAGCAACAATATCATCTGCAACA 979
Db 272 ProThrLeuThrIleSerGlyAsnLysSerLeuThrPheAlaGluAsnSerSerValThr 291
QY 980 GCGGAGGAGCTATCTATGTGAAAGCTCGAAGTGGTTCGCGAGGACTTACCTATTTC 1039
Db 292 GlnGlyGlyAlaIleCysAlaHisGlyLeuAspLeuSerAlaAlaGlyProThrLeuPhe 311
QY 1040 AGTAGAAATAGTGTCAATGGAGGTACAGCTCTCTAAAGGTGGAGCCATAGCTATCGAAGT 1099
Db 312 SerAsnAsnArgCysGlyAsnThrAlaAlaGlyLysGlyGlyAlaIleAlaIleAlaAsp 331
QY 1100 AGTGGGAATTCAGTTTATCCCGCATGATGTGTCATTTGTTTTCAGGGAATCAGTCT 1159
Db 332 SerGlySerLeuSerLeuSerAlaAsnGlnGlyAspIleThrPheLeuGlyAsnThrLeu 351
QY 1160 ACTTCTACTACT---CCTGGGACGATAGAACTAGTATGCTATGAGTACGAGTCAAGTCAAG 1216
Db 352 ThrSerThrSerAlaProThrSerThrArgAsnAlaIleTyrrLeuGlySerSerAlaLys 371
QY 1217 ATGACAGCTTTCTGCTGCTGCTAGAGCATCTACTTCTATGATGCCATA----- 1270
Db 372 IleThrAsnLeuArgAlaAlaGlnGlyGlnSerIleTyrrPheTyrrAspProIleAlaSer 391
QY 1271 ---ACTACAGATCATCCACAACAGTTACAGATGTCTTAAAGTTTAAGTACGACTCGGCA 1327
Db 392 AsnThrThrGlyAlaSer-----AspValLeuThrIleAsnGlnProAspSer 407
QY 1328 GATTCGTGCACTAATATACAGGGACATCATCTTTCACAGGAGAAAGTTATCAGAGACA 1387
Db 408 AsnSerProLeuAspTyrrSerGlyThrIleValPheSerGlyGluLysLeuSerAlaAsp 427
QY 1388 GAGGCGCGCAGATCTTAAATACTTACTTTCGAAAGTACTACAGCTGTAACTTTTCAGGA 1447
Db 428 GluAlaLysAlaAlaAspAsnPheThrSerIleLeuLysGlnProLeuAlaLeuAlaSer 447
QY 1448 GGTACTCTATCTTTTAAACATCGAGTGACTCTGCGAGCTCAGGCAATTCACCAACAGGCA 1507
Db 448 GlyThrLeuAlaLeuLysGlyAsnValGluLeuAspValAsnGlyPheThrGlnThrGlu 467
QY 1508 GATTCGTGCTCGAAATGAGCTAGGAATCTCTAGAACCTGCTGATAGTACCTAGTACCAT 1567
Db 468 GlySerThrLeuLeuMetGlnProGlyThrLysLeuLys---AlaAspThrGluAlaIle 486
QY 1568 AAC-----AATTTGCTCATTAACATCAGTTCTATAGACGGTGCAGAAAGGCAAAATA 1621
Db 487 SerLeuThrLysLeuValValAspLeuSerAlaLeuGluGlyAsnLysSerValSerIle 506
QY 1622 GAAACAAGCTACGTCAAAAATCTGACTTTATCTGGAACCATCCTTTTATTGGACCCG 1681
Db 507 GluThrAlaGlyAlaAsnLysThrIleThrLeuThrSerProLeuValPheGlnAspSer 526
QY 1682 ACGGCGACGCTTTTATGAAATCATAGTTTAAAGATCTCTCAGTCTCTCAGCATCTTAGAG 1741
Db 527 SerGlyAsnPheTyrrGluSerHisThrIleAsnGlnAlaPheThrGlnProLeuValVal 546
QY 1742 CTCAAAGCTTCT-----GGAACTGTAACTAACAGCAGCCGCA 1774
Db 547 PheThrAlaAlaThrAlaAlaSerAspIleTyrrIleAspAlaLeuLeuThrSerProVal 566

```


QY 1775 GTGACTCCAGATCTATATGGGTGAGAAATTCATTACGGCTATCAGGAACTTGGGG 1834
 Db |||||:|||||
 567 GlnThrProGluPro-----HisTyrGlyTyrGlnGlyHisTrpGlu 580
 QY 1835 CCAATTGTTGGGGGACAGGGGCTTACGACTGCAACCTTCAACTGGACTAAACCTGGC 1894
 Db |||||:|||||
 581 AlaThrTrpAlaAspThrSerThrAlaLysSerGlyThrMetThrTrpValThrThrGly 600
 QY 1895 TATATTCCTAATCCGAGCGTATCGGCTCTTAGTCCCTAAATAGCTTATGGAAATGCAATTT 1954
 Db |||||:|||||
 601 TyrAsnProAsnProGluArgArgAlaSerValValProAspSerSerLeuTrpAlaSerPhe 620
 QY 1955 ATAGATTATAGCTCTCTCCATTTATCTATGAGACTGCAACGAAAGGTTGACGAGAGAC 2014
 Db |||||:|||||
 621 ThrAspIleArgThrLeuGlnGlnIleMetThrSerGlnAlaAsnSerIleTyrGlnGln 640
 QY 2015 CGTCTTTTGGTGCTCGAATTATCTAATCTCTCCATAAGGATAGTACAAACACACGA 2074
 Db |||||:|||||
 641 ArgGlyLeuTrpAlaSerGlyThrAlaAsnPhePheHisLysAspLysSerGlyThrAsn 660
 QY 2075 CGCGGTTTCGCATTTAGTGGCGGTATGTCATAGGAGGAAACCTACATCTTGTTCA 2134
 Db |||||:|||||
 661 GlnAlaPheArgHisLysSerTyrGlyTyrIleValGlyGlySerAlaGluAspPheSer 680
 QY 2135 GATAAGATTCTTAGTCTGCTCATTTCTCAGCTCTTTGGAGAGATAGAGACTACTTGTGA 2194
 Db |||||:|||||
 681 GluAsnIlePheSerValAlaPheCysGlnLeuPheGlyLysAspLysLeuPheIle 700
 QY 2195 GCTAAGAACTCAAGGTACAGTCTACGAGGAACTCTCTATTACGACACACCAACCACTAT 2254
 Db |||||:|||||
 701 ValGluAsnThrSerHisAsnTyrLeuAlaSerLeuTyrLeuGlnHisArgAlaPheLeu 720
 QY 2255 ATCTCTCTCTTGCMAACTACGCGCTTGT-----TCGTTGCTTATGTTCTCTACAGAG 2308
 Db |||||:|||||
 721 GlyGlyLeuPro-----MetProSerPheGlySerIleThrAspMetLeuLysAsp 737
 QY 2309 ATTCTGTTCTCTTTTCAGGAACCTTAGTCTACCCATACCGATACCGATCTGAAACC 2368
 Db |||||:|||||
 738 IleProLeuIleLeuAsnAlaGlnLeuSerTyrSerTyrThrLysAsnAspMetAspThr 757
 QY 2369 AAGTATACAAATATCTACTGTAAAGAAAGCTGGGGAATGATAGTTTCGCTTTAGAA 2428
 Db |||||:|||||
 758 ArgTyrThrSerTyrProGluAlaGlnGlySerTyrThrAsnAsnSerGlyAlaLeuGlu 777
 QY 2429 TTCGGTGAAGCTCCGATTGCTTA-----GATGAAGAGTCTCTATTGAGCAGTACATG 2485
 Db |||||:|||||
 778 LeuGlyGlySerLeuAlaLeuTyrLeuProLysGluAlaProPhePheGlnGlyTyrPhe 797
 QY 2486 CCCTTCATGAAATTCAGTTTCTCTATGCACATCAGGAGGTTTAAAGAACAGGGAACA 2545
 Db |||||:|||||
 798 ProPheLeuLysPheGlnAlaValTyrSerArgGlnGlnAsnPheLysGluSerGlyAla 817
 QY 2546 GAAGCTCGTGAATTTGGAAGTAGCGCTTGTGAATCTTGCTTACCTATCGGGATCCGA 2605
 Db |||||:|||||
 818 GluAlaArgAlaPheAspGlyAspLeuValAsnCysSerIleProValGlyIleArg 837
 QY 2606 TTTGATAGGAATCAGATCGCAAGATGCAACGTAACATCTCACTCTTGTTTATCTGTG 2665
 Db |||||:|||||
 838 LeuGluLysIleSerGluAspGlyLysAsnAsnPheGluIleSerLeuAlaTyrIleGly 857
 QY 2666 GATCTTGTTCGAGTAAACCCGACTGTACGACAACTCGCAATAGCGGTGATCTTGG 2725
 Db |||||:|||||
 858 AspValTyrArgLysAsnProArgSerArgThrSerLeuMetValSerGlyAlaSerTrp 877
 QY 2726 AAAACCTTCGGTACGAATTTGCAAGACAAAGCTTTAGTCTTCGTCGAGGGAACCATTTT 2785
 Db |||||:|||||
 878 ThrSerLeuCysLysAsnLeuAlaArgGlnAlaPheLeuAlaSerAlaGlySerHisLeu 897
 QY 2786 TGCCTTAATCAAAATTTGAAGCCTTTAGCCCAATTTCTTTTGAATTCGCTGGGTCATCT 2845
 Db |||||:|||||
 898 ThrLeuSerProHisValGluLeuSerGlyGluAlaAlaTyrGluLeuArgGlySerAla 917

QY 2846 CGCAATTACAATCTAGACTTAGAGCAAAATACCAATTC 2884
 Db |||||:|||||
 918 HisIleTyrAsnValAspCysGlyLeuArgTyrSerPhe 930
 RESULT 5
 ID PM11 CHLPN STANDARD; PRT; 928 AA.
 AC 086164; Q9K299;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable outer membrane protein pmpl1 precursor (Polymorphic membrane
 DE protein 11) (outer membrane protein 4).
 GN PMP11 OR OMP4 OR CPN0449 OR CP0302.
 OS Chlamydia pneumoniae (Chlamydia phila pneumoniae).
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phila.
 OX NCBI_TaxID=83558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CWL029/VR-1310;
 RX MEDLINE=99081766; PubMed=9864239;
 RA Knudsen K., Madsen A.S., Mygind P., Christiansen G., Birkelund S.;
 RT "Identification of two novel genes encoding 97- to 99-kilodalton outer
 RT membrane proteins of Chlamydia pneumoniae.";
 RL Infect. Immun. 67:375-383(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VR1310;
 RX MEDLINE=20007584; PubMed=10539856;
 RA Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
 RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;
 RT "Molecular biology of Chlamydia pneumoniae surface proteins and their
 RT role in immunopathogenicity.";
 RL Am. Heart J. 138:S491-S495(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CWL029;
 RX MEDLINE=99206606; PubMed=10192388;
 RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
 RL Nat. Genet. 21:385-389(1999).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AR39;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
 RT pneumoniae AR39.";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J138;
 RX MEDLINE=20330349; PubMed=10871362;
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 RT from Japan and CWL029 from USA.";
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
 CC (POTENTIAL).
 CC -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)


```
Db 562 TyrGlyPheGlnGlyGlnTrpAsn---ValaenTrrThrThrAspThrAlaThrAsnThr 580
QY 1868 -----GCAACCTTCACTGGAGTAAACTGGCTATATCTTAATCCGAGCGTATCGGC 1921
Db 581 LysGluAlaThrAlaThrTrpThrLysThrGlyPheValProSerProGluArgLysSer 600
QY 1922 TCTTTAGTCCCTTAATAGCTTATGGAATGATGATTAATAGATATATAGCTCTCTCCATATCTT 1981
Db 601 AlaLeuValCysAsnThrLeuTrpGlyValPheThrAspIleArgSerLeuGlnGlnLeu 620
QY 1982 ATGGAGACTGCACAAAGAGGTTGCAGGAGACCGCTGCTTTTGGTGCTGCTGATATCT 2041
Db 621 ValGluIleGlyAlaThrGlyMetGluHisLysGlnGlyPheTrpValSerSerMetThr 640
QY 2042 AACTTCTTCCATAAGATAGTACAAAACACACGCGGTTTCGCCATTTGAGTGGCGGT 2101
Db 641 AsnPheLeuHisLysThrGlyAspGluAsnArgLysGlyPheArgHisThrSerGlyGly 660
QY 2102 TATGTATAGGAGGAAACCTACATCTTGTTCAGATAAGATCTTAGTGTGCTGCTATGTTGT 2161
Db 661 TyrValIleGlyGlySerAlaHisThrProLysAspLeuPheThrPheAlaPheCys 680
QY 2162 CAGCTCTTTGGAAGAGATAGAGACTACTTTGTAGTAAAGATCAAGGTACAGTCTACGGA 2221
Db 681 HisLeuPheAlaArgAspLysAspCysPheIleAlaHisAsnAsnSerArgThrTyrgly 700
QY 2222 GGAACCTCTATTACACGACACACAAACCC-----TATATCTCTCTT--- 2263
Db 701 GlyThrLeuPhePheIleHisSerHisThrLeuGlnProGlnAsnTyLeuArgLeuGly 720
QY 2264 CCTTGCAAACTACGGCCTTGTTCGTCTTATGTTCTCTACAGATTCCTGTTCTCTTT 2323
Db 721 ArgAlaLysPheSerGluSerAlaIleGluLysPheProArgGluIleProLeuAlaLeu 740
QY 2324 TCAGAAACCTTAGTACACCATACGGATACAGTCTGAAACCAACGATATACACATAT 2383
Db 741 AspValGlnValSerPheSerHisSerAspAsnArgMetGluThrHisTyThrSerLeu 760
QY 2384 CCTACTGTTAAAGAGTGGGGAGTATAGTATTCGCTTTAGATTCGGTGGGAGAGCT 2443
Db 761 ProGluSerGluGlySerTrpSerAsnGluCysIleAlaGlyIleGlyLeuAspLeu 780
QY 2444 CCGATTTTCTCTTA---GATGAAAGTCTCTTATTGAGCAGTACATGCCCTTCATGAAATTG 2500
Db 781 ProPheValLeuSerAsnProHisProLeuPheLysThrPheIleProGlnMetLysVal 800
QY 2501 CAGTTTGTCTATGCACATCAGAAAGGTTTAAAGAACAGGGAACAGAGCTCGTGAATTT 2560
Db 801 GluMetValTyrrValSerGlnAsnSerPhePheGluSerSerAspGlyArgGlyPhe 820
QY 2561 GGAAGTAGCCCTCTGTTGAATCTTCCTATCGGATCCGATTCGATTTGATGAAGATCA 2620
Db 821 SerIleGlyArgLeuLeuAsnLeuSerIleProValGlyAlaLysPhe---ValGlnGly 839
QY 2621 GACTGCCAAGATGCA---ACGTACAATCTAATCTCTGTTGTTACTGTGATCTTGTTCGT 2677
Db 840 AspIleGlyAspSerTyThrTyrrAspLeuSerGlyPhePheValSerAspValTyrrArg 859
QY 2678 AGTAACCCGACTGTACGACACACTGCGAATAGCGGTGATCTTGTGAAACCTTCGGT 2737
Db 860 AsnAsnProGlnSerThrAlaThrLeuValMetSerProAspSerTrpLysIleArgGly 879
QY 2738 ACGAATTTGGACAGACAGCTTTAGTCTCTGTCGAGGGAACCAATTTTGTCTTAATCTCA 2797
Db 880 GlyAsnLeuSerArgGlnAlaPheLeuLeuArgGlySerAsnAsnTyrrValTyrrAsnSer 899
QY 2798 AATTGTAAGCCTTTAGCCAAATTTCTTTTCAATTGCGTGGGTCTATCTCGCAATTAACAT 2857
Db 900 AsnCysGluLeuPheGlyHisTyrrAlaMetGluLeuArgGlySerArgAsnTyrrAsn 919
QY 2858 GTAGACTTAGGACCAAAATACCAATTC 2884
```

```
Db 920 ValaspValGlyThrLysLeuArgPhe 928
RESULT 6
PMP2_CHLPN
ID PMP2_CHLPN STANDARD; PRT; 841 AA.
AC Q9Z3A1; Q9RB73;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable outer membrane protein pmp2 precursor (Polymorphic membrane
protein 2) (Outer membrane protein 7).
GN PMP2 OR OMP7 OR CPN0013 OR CP0761.
OS Chlamydia pneumoniae (Chlamydia phylum pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phylum.
OX NCBI_Taxid=83558;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=VR1310;
RX MEDLINE=20007584; PubMed=10539856;
RA Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT "Molecular biology of Chlamydia pneumoniae surface proteins and their
role in immunopathogenicity.";
RL Am. Heart J. 138:S491-S495(1999).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [3]
SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., Bass S.,
White O., Hickey E.K., Peterson J., Utterback R., Berry K., Bass S.,
Lihner K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
pneumoniae A839.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [4]
SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
Shiba T., Ishii K., Hattori M., Kohara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
(POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC -!- CAUTION: REF.4 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
FRAMESHIFT IN POSITION 673.
-----
THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.ebi.ac.uk/ebis/
or send an email to license@ebi.ac.uk).
-----
EMBL; AJ133035; CAB37083.1; -
EMBL; AE001586; AAD18172.1; -
EMBL; AF002235; AAF38561.1; -
EMBL; AP002545; BAA98223.1; ALT_FRAME.
PIR; E72130; E72130.
PHCI-2DPAGE; Q9Z3A1; -.
```

TIGR; CP0761;	-
DR InterPro; IPR003368; Chlamydia_PMP.	
DR Pfam; PF02415; DUF145; 1.	
DR TIGRFAMs; TIGR01376; POMP repeat; 3.	
KW Outer membrane; Signal; Multigene family; Complete proteome.	
FT SIGNAL	1 24 POTENTIAL.
FT CHAIN	25 841 PROBABLE OUTER MEMBRANE PROTEIN PMP2.
FT CONFLICT	784 784 N -> S (IN REF. 4).
SQ SEQUENCE	841 AA; 89600 MW; 9064D6D0678D24C CRC64;
Alignment Scores:	
Pred. No.:	1.07e-87 Length: 841
Score:	1580.50 Matches: 375
Percent Similarity:	52.54% Conservative: 121
Best Local Similarity:	39.72% Mismatches: 306
Query Match:	29.22% Indels: 142
DB:	1 Gaps: 19
US-09-428-122-1 (1-3000) x PMP2_CHLPN (1-841)	
QY 134 ACAITTTGCTATTTCCCTTTCCTATGATGTCTACCGAGACAGATTGGATTCAAGTCGC	193
Db 15 ThrLeuSerMetSerAsnLeuLeuGlyAlaAlaThrThrGluGluSerAlaSerAsn	34
QY 194 AGTTTCOGATGGG--AATAAAAAATGTAATTTTTTCAGTTCGTGAGAGTCAGGAAGATGCT	250
Db 35 SerPheAspGlyThrThrSerThrThrSerPheSerSerLysThrSerAlaThrAsp	54
QY 251 GGAACTACTACCTATTAAAGGAATGCTCACTTAGAAAATATTCTTGGACACGCCACA	310
Db 55 GlyThrAsnTyrValPheLysAspSerValIleGluAsnValProLysThrGlyGlu	74
QY 311 GCAATCACAAAAGCTGTTTTTAAACAAC-----ACTAAGGGCGGATTTCAGCTTTCACAGT	364
Db 75 ThrGlnSerThrSerCysPheLysAsnAspAlaAlaGlyAspLeuAsnPheLeuGly	94
QY 365 AACGGGAACCTCTCATVNTGTTCCAAACGGTGAGTCAGGGACGTGTAGACGGGCTCGTGT	424
Db 95 GlyGlyPheSerPheThrPheSerAsnIleAspAlaThrThrAlaSerGlyAlaAlaIle	114
QY 425 AACACACAGCGTGTAGATAAAACTACCAGTTTTATAGGTTTTCTTCGCTATCTTTATT	484
Db 115 GlySerGluAlaAlaAsnLysThrValThrLeuSerGlyPheSerAlaLeuSerPheLeu	134
QY 485 CGCTCTCTCGGAAGTTCGATAACTACCGCAAAGGAGCGGTAGCTGCTCTACGGGTAGC	544
Db 135 LysSerProAlaSerThrValThrAsnGlyLeuGlyAlaIleAsnValLys---GlyAsn	153
QY 545 TTGAGTTTGACAAAATGTCAGTTTGCTCTTCAGCAAAAACCTTTTCACGGATATGCG	604
Db 154 LeuSerLeuLeuAspAsnAspLysValLeuIleGlnAspAsnPheSerThrGlyAspGly	173
QY 605 GTGTGTATCACCGCAAAAACCTCTTTCATTAAACAGGAGCTACAATGTCTGTTTCTT	664
Db 174 GlyAlaIleAsnCys----	178
QY 665 GAAATACCTCTCAAAGAAAGGAGGCCATTACAGCTTCGATGCCCTTACCATTACT	724
Db 178 -----	178
QY 725 GGAACCAAGGGAAGTCTCTTTTCTGACAATACTTCTTCGGATTCTCGAGCTGCAATT	784
Db 178 -----	178
QY 785 TTTCAGAAGCCTCGGTGACTATTTCTAATAATGCTPAAAGTTTCCTTTATTGACAATAAG	844
Db 179 -----AlaGlySerLeuLysIleAlaAsnAsnLysSerLeuSerPheIle	193
QY 845 GTCACAGGAGCGAGCTCCTCAACACGGGGGATATGTCAGGAGGTCTATCTGTGCTTAT	904
Db 193 -----	193
QY 905 AAAACTAGTACAGATACTAAGTTCACCCCTCACTGGAAATCAGATGTTACTTCTTCAGAAC	964

Db 524 ValArgSerPheGlnAsnPheIleGluLeuGlyThrGluGlyAlaProTyrGluLysArg 543
 QY 2021 TTTTCGTCGTCGATTACTAATCTTCCATAAGTAGTACAAAACACAGCGGG 2080
 Db 544 PheIlePValAlaGlyIleSerAsnValLeuHisArgSerGlyArgGluAsnGlnArgLys 563
 QY 2081 TTTCCGCTTTGAGTGGCGGTATGTCTATGATAGGAGAACCTACATATGTTTCAGATAAG 2140
 Db 564 PheArgHisValSerGlyGlyAlaValValGlyAlaSerThrArgMetProGlyGlyAsp 583
 QY 2141 ATTCTTAGTCTGCATTTTGTCTAGCTTTTGGAGAGATAGAGACTTCTTGTAGCTAAG 2200
 Db 584 ThrLeuSerLeuGlyPheAlaGlnLeuPheAlaArgAspLysAspTyrPheMetAsnThr 603
 QY 2201 AATCAAGGTACAGTCTACGGAGAACTCTCTATTACCGACACGAAACCTATATCTCT 2260
 Db 604 AsnPheAlaLysThrTyrAlaGlySerLeuArgLeuGlnHisAspAlaSerLeuTyrSer 623
 QY 2261 CTTCTCTGCTG-----AACTACCGCTTGTCTGTCTTATCTTCTCT 2302
 Db 624 ValValSerIleLeuLeuGlyGluGlyLeuArgGluIleLeuLeuProTyrValSer 643
 QY 2303 ACAGAGATTCCTGTTCTCTTTTCAGAAACCTTACCTACCCATACGGAATAACGATCTG 2362
 Db 644 LysThrLeuProCysSerPheTyrGlyGlnLeuSerTyrGlyHisThrAspHisArgMet 663
 QY 2363 AAACCAAGTAT-----ACAACATATCTCTCTGTTAAAGGAGCTGGGG 2407
 Db 664 LysThrGluSerLeuProProProProProProProProProProProProProProPro 683
 QY 2408 AATGATGTTTTCGTTTGAATTCGTTGGAAGAGCTCCGATTTCTTGTAGTGAAGTCT 2467
 Db 684 GlyTyrValTyrAlaGlyGluLeuGlyThrArg-----ValAlaValGluAsnThrSer 701
 QY 2468 -----CTATTGACAGTACATGCGCTTCATGAATTCGATTTGCTATGACAT 2518
 Db 702 GlyArgGlyPhePheGlnGluTyrThrProPheValLysValGlnAlaValTyrAlaArg 721
 QY 2519 CAGGAAGTTTAAAGACAGGACAGAGCTCGTGAATTTGGAAGTAGCGCTCTGTG 2578
 Db 722 GlnAspSerPheValGluLeuGlyAlaIleSerArgAspPheSerAspSerHisLeuTyr 741
 QY 2579 AATCTTGCTTACCTTACCGATCCGATTCGATTTGATAGGAATCAGACTGCCAAGATGCAACG 2638
 Db 742 AsnLeuAlaIleProLeuGlyIleLysLeuGluLysArg-----PheAlaGluGlnTyr 759
 QY 2639 TACATCTAATCTCTGTTTACTGATCTGATCTGTTCTGTTAGTAAACCGACTGACGACA 2698
 Db 760 TyrHisValValAlaMetTyrSerProAspValCysArgSerAsnProLysCysThrThr 779
 QY 2699 ACATGCGAATTAGCGGTGATTCTTGGAAACCTTCGCTAGCAATTTGCGCAAGCAAGCT 2758
 Db 780 ThrLeuLeuSerAsnGlnGlySerThrLysThrLysGlySerAsnLeuAlaArgGlnAla 799
 QY 2759 TTATGCTCTGTCGAGGAACCATTTTGTCTTTAACTCAATTTTGAAGCCCTTAGCCAA 2818
 Db 800 GlyIleValGlnAlaSerGlyPheArgSerLeuGlyAlaAlaAlaGluLeuPheGlyAsn 819
 QY 2819 TTTTCTTTTGAATTCGCGGTCTATCTCGCAATTAATGATAGACTTAGGAGCAAAATAC 2878
 Db 820 PheGlyPheGluTyrPargGlySerSerArgSerTyrAsnValAspAlaGlySerLysIle 839
 QY 2879 CAATTC 2884
 Db 840 LysPhe 841

RESULT 7

PMP6_CHLPN

ID -PMP6_CHLPN STANDARD; PRT; 1276 AA.
 AC Q92899; OSJRW2;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Probable outer membrane protein pmp6 precursor (Polymorphic membrane
 protein 6).
 GN PMP6 OR CFN0444 OR CP0309.
 OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
 OX NCBI_TaxID=83558;
 RN [1] _SEQUENCE FROM N.A.
 RP STRAIN=CNL029;
 RC MEDLINE=9920606; PubMed=10192388;
 RX Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
 RL Nat. Genet. 21:385-389(1999).
 RN [2] _SEQUENCE FROM N.A.
 RP STRAIN=AR39;
 RC MEDLINE=20150255; PubMed=10684935;
 RX Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
 pneumoniae AR39";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 RN [3] _SEQUENCE FROM N.A.
 RP STRAIN=J138;
 RC MEDLINE=20330349; PubMed=10871362;
 RX Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 from Japan and CWL029 from USA";
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -|- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
 (POTENTIAL).
 CC -|- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 DR EMBL; AE001627; AAD18588.1; -;
 DR EMBL; AE002193; AAF38166.1; -;
 DR EMBL; AP002546; BAA98652.1; -;
 DR PIR; B72078; B72078.
 DR PIR; B86546; B86546.
 DR PIR; C81591; C81591.
 DR PHCI-2DPAGE; Q92899; -;
 DR TIGR; CP0309; -;
 DR InterPro; IPR006315; Autotransport.
 DR InterPro; IPR003368; Chlamydia_PMP.
 DR Pfam; PF02415; DUF145; 2.
 DR TIGRFAMs; TIGR01414; autotrans_bar1; 1.
 DR TIGRFAMs; TIGR01376; POMP repeat; 14.
 DR Outer membrane; Signal; Multigene family; Complete proteome.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 1276 PROBABLE OUTER MEMBRANE PROTEIN PMP6.
 FT CONFLICT 421 421 E -> ECGGAILAFIDSGVSKTGLSIANNQEVSLTNSA
 FT FTLTGSGTGTFTSTNTAKTGALYSKGNNSISGNTLLFSG
 FT NKATGPSNSANOE (IN REF. 1).
 SQ SEQUENCE 1276 AA; 132127 MW; C97FA0662C88E0B CRC64;
 Alignment Scores:
 Pred. No.: 1.82e-79 Length: 1276
 Score: 1444.00 Matches: 382
 Percent Similarity: 41.62% Conservative: 157

Best Local Similarity: 29.50% Mismatches: 370			
Query Match: 26.70% Indels: 386			
DB: 1 Gaps: 26			
US-09-428-122-1 (1-3000) x PMP6_CHLPN (1-1276)			
QY	101	ATGAAGTCTCTTCCCAAGTTTGATTTTCTACATTTGCTATTTTCCCTTTG---TCT	157
Db	1	MetLysThrLeuProThrLeuLeuThrSerSerAlaLeuValPheSerLeuHisPro	20
QY	158	ATGATTGCTACCGAGACAGTTTGGATTCAAGTCGAGTTTCGATGGGAATAAAATGCT	217
Db	21	LeuMetAlaAlaAsnThrAspLeuSerSerAspAsnThrGluAsnGlySerSerGly	40
QY	218	AAT-----TTTTCACTTCGTGAGTCAGGAGATGCTGGAACCTACCTACCTATTAAAG	271
Db	41	SerAlaAlaPheThrAlaLysGluThrSerAspAlaSerGlyThrThrThrThrLeuThr	60
QY	272	GGAAATGTCACCTAGAAATATTCCTGGAAACAGGCACAGCAATCACA-----AAA	322
Db	61	SerAspValSerIleThrAsnVal-----SerAlaIleThrProAlaAspLys	76
QY	323	AGCTGTTTTAAACAACACTAAGCGGATTTCACTTTCCACAGGTAACGGGAACCTCTCTATTG	382
Db	77	SerCysPheThrAsnThrGlyGlyAlaLeuSerPheValGlyAlaAspHisSerLeuVal	96
QY	383	TTCCAAACGGTGAGTACGAGGACTGTAGCAGGGGCTGCTGTTAAACAGCAGCGGTGAT	442
Db	97	LeuGlnThrIle---AlaLeuThrHisAspGlyAlaAlaIleAsnAsnThr-----Asn	113
QY	443	AAATCTACCACGPTTATAGGTTTTCTTCGCTATCTTTATTGCGTCTCTGGAAGTCG	502
Db	114	ThrAlaLeuSerPheSerGlyPheSerSerLeuLeuIleAspSerAlaProAlaThrGly	133
QY	503	ATAACTACCGCAAGAGCGGTAGCTGCTCTACG-----GCTAGCTTGAGT	550
Db	134	ThrSerGlyGlyLysGlyAlaIle---CysValThrAsnThrGluGlyGlyThrAlaThr	152
QY	551	TTGACAAAAGTCACTGTTGCTCTTCAGCAAAAACCTTTTCAACGGATTAATGGCGTGCT	610
Db	153	PheThrAspAsnAlaSerValThrLeuGlnLysAsnThrSerGluLysAspGlyAlaAla	172
QY	611	ATCACCGCAAAACTCTTTCATTAAACAGGACTACAAATGTCAGCTCTGTTTCTGAAAT	670
Db	173	ValSerAlaLysSerIleAspLeuAlaLysThrThrAlaAlaLeuLeuAspGlnAsn	192
QY	671	ACCTCTCAAGAAAGCGGAGCCATTCACTTCCGATCCCTTACCATTACTGGAAC	730
Db	193	ThrSerThrLysAsnGlyGlyAlaLeuCysSerThrAlaAsnThrThrValGlnGlyAsn	212
QY	731	CAAGGGAGTCTCTTTTCTGCAATACTTCTCGGATTCGGAGCTGCAATTTT---	787
Db	213	SerGlyThrValThrPheSerSerAsnThrAlaThrAspLysGlyGlyIleThrSer	232
QY	787	-----	787
Db	233	LysGluLysAspSerThrLeuAspAlaAsnThrGlyValValThrPheLysSerAsnThr	252
QY	788	-----ACAGAGCCTCGGTGACTATTCTTAATAATGCTAAA	823
Db	253	AlaLysThrGlyGlyAlaThrSerSerAspAsnLeuAlaLeuThrGlyAsnThrGln	272
QY	824	GTTTCTTTTATTGACAAATAGGTCACAGGA---GCGAGTCCTCAACAACGGGGATATG	880
Db	273	ValLeuPheGlnGluAsnLysThrThrGlySerAlaAlaGlnAlaAsnProGluGly	292
QY	881	TCAGAGGAGTCTCTGCTGCTAT---AAAACCTAGTACAGAT-----	919
Db	293	CysGlyGlyAlaIleCysCysThrLeuAlaThrAlaThrAspLysThrGlyLeuAlaIle	312
QY	919	-----	919
Db	313	SerGlnAsnGlnGluMetSerPheThrSerAsnThrThrThrAlaAsnGlyGlyAlaIle	332
QY	920	-----ACTAAGTCCACCTCAGTGAATCAGATGTTACTTCTTCCAGCAATACATCG	973
Db	333	TyrAlaThrLysCysThrLeuAspGlyAsnThrThrLeuThrPheAspGlnAsnThrAla	352
QY	974	ACA-----	976
Db	353	ThrAlaGlyCysGlyGlyAlaIleThrThrGluThrGluAspPheSerLeuLysGlySer	372
QY	976	-----	976
Db	373	ThrGlyThrValThrPheSerThrAsnThrAlaLysThrGlyGlyAlaLeuThrSerLys	392
QY	976	-----	976
Db	393	GlyAsnSerSerLeuThrGlyAsnThrAsnLeuLeuPheSerGlyAsnLysAlaThrGly	412
QY	976	-----	976
Db	413	ProSerAsnSerSerAlaAsnGlnGluGlyCysGlyGlyAlaIleLeuSerPheLeuGlu	432
QY	976	-----	976
Db	433	SerAlaSerValSerThrLysLysGlyLeuTrpIleGluAspAsnGluAsnValSerLeu	452
QY	976	-----	976
Db	453	SerGlyAsnThrAlaThrValSerGlyGlyAlaIleThrAlaThrLysCysAlaLeuHis	472
QY	976	-----	976
Db	473	GlyAsnThrThrLeuThrPheAspGlyAsnThrAlaGluThrAlaGlyGlyAlaIleTyr	492
QY	976	-----	976
Db	493	ThrGluThrGluAspPheThrLeuThrGlySerThrGlyThrValThrPheSerThrAsn	512
QY	976	-----	976
Db	513	ThrAlaLysThrAlaGlyAlaLeuHisThrLysGlyAsnThrSerPheThrLysAsnLys	532
QY	976	-----	976
Db	533	AlaLeuValPheSerGlyAsnSerAlaThrAlaThrAlaThrThrThrAspGlnGlu	552
QY	976	-----	976
Db	553	GlyCysGlyGlyAlaIleLeuCysAsnIleSerGluSerAspIleAlaThrLysSerLeu	572
QY	976	-----	976
Db	573	ThrLeuThrGluAsnGluSerLeuSerPheIleAsnAsnThrAlaLysArgSerGlyGly	592
QY	976	-----	976
Db	593	GlyIleThrAlaProLysCysValIleSerGlySerGluSerIleAsnPheAspGlyAsn	612
QY	977	-----ACAGCGGAGGAGCTATCTATGTGAAAAAGCTCGAATCGCTTCCGAGGA	1027
Db	613	ThrAlaGluThrSerGlyGlyAlaIleThrSerLysAsnLeuSerIleThrAlaAsnGly	632
QY	1028	CTTACCTTATTCAGTAGAATAAGTGTCAATGGAGGTACAGCTCCCTAAAGGTGGAGCCATA	1087
Db	633	ProValSerPheThrAsnAsnSer-----GlyGly-----LysGlyGlyAlaIle	647
QY	1088	GCTATCGAAGATAGTGGGAATTGAGTTTATCCCGGATAGTGTGACATGCTCTTTTA	1147
Db	648	TyrIleAlaAspSerGlyGluLeuSerLeuGluAlaIleAspGlyAspIleThrPheSer	667
QY	1148	GGGAATACAGTCTTCTACTACTCTCTGGGAGCAATAGAGTAGTAGTCTAGCTAGCAACG	1207
Db	668	GlyAsnArgAlaThrGluGlyThrSerThrProAsn-----SerIleHisLeuGlyAla	685

RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
 RL Nat. Genet. 21:385-389(1999).
 [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=AR39;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.W.;
 RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
 RT pneumoniae AR39.";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 [4]
 RN SEQUENCE FROM N.A.
 RC STRAIN=J138;
 RX MEDLINE=20330349; PubMed=10871362;
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 RT from Japan and CWL029 from USA";
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -/- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
 CC -/- (POTENTIAL).
 CC -/- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR ENBL; AJ133035; CAB37075.1; -
 DR ENBL; AE001585; AAD18163.1; -
 DR ENBL; AE002237; AAF38570.1; -
 DR ENBL; AP002545; BAA98215.1; -
 DR PIR; B72131; B72131.
 DR PIR; B86491; E86491.
 DR TIGR; CP0770; -.
 DR InterPro; IPR003368; Chlamydia_PMP.
 DR Pfam; PF02415; DUF145; 1.
 DR TIGRFAMS; TIGR01376; POMP repeat; 5.
 KW Outer membrane; Signal; Multigene family; Complete proteome.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 922 PROBABLE OUTER MEMBRANE PROTEIN PMP1.
 FT CONFLICT 14 14 F -> L (IN REF. 1).
 FT CONFLICT 375 375 Y -> C (IN REF. 1).
 FT CONFLICT 606 606 D -> N (IN REF. 1).
 SQ SEQUENCE 922 AA; 100457 MW; DFF2AB6333AB031C CRC64;
 Alignment Scores:
 Pred. No.: 2.12e-79 Length: 922
 Score: 1442.50 Matches: 351
 Percent Similarity: 53.31% Conservative: 156
 Best Local Similarity: 36.91% Mismatches: 381
 Query Match: 26.67% Indels: 63
 DB: 1 Gaps: 26
 US-09-428-122-1 (1-3000) x PMP1_CHLPN (1-922)
 QY 113 TTCCCC--AAGTTGTGATTTCTACATTTGCTATTTTCCCTTTGCTATGATTGCTACC 169
 DB 8 PheProLeuValPheSerPheThrLeuLeuSerValPheAspThrSerLeuSerAlaThr 27
 QY 170 GAGACAGTTTGGATTCAAGTCGCGAGTTTCGATGGGAATAAAATGTAATTTTCAGTT 229
 DB 28 ThrIleSerLeuThrProGluAspSerPheHisGlyAspSerGlnAsn-----AlaGlu 45

QY 230 CGTGAGAGTCAGAGAGATCTCGAACTACCTACCTATTATTAAGGGAATCTCACTCTAGAA 289
 DB 46 ArgSerTyrAsnValGlnAlaGlyAspValTyrSerLeuThrGlyAspValSerIleSer 65
 QY 290 AATATTCTCGGAACAGGCACAGCAATCACAAAAGCTGTTTAAACAACAACAAAGGCGAT 349
 DB 66 AsnVal-----AspAsnSerAlaLeuAsnLysAlaCysPheAsnValThrSerGlySer 83
 QY 350 TTGACTTTCCAGGTAAACGGGAACCTCTATTGTTCCAAACGGTGGATCAGGAGACTGTA 409
 DB 84 ValThrPheAlaGlyAsnHisGlyLeuTyrPheAsnAsnIleSerSerGlyThrThr 103
 QY 410 GCAGGGCTGCTGTTTAAACAGCAGCGTGGTAGATAAATCTACCACG---TTTATAGGTTT 466
 DB 104 LysGluGlyAlaValLeuCysCysGlnAspProGlnAlaThrAlaArgPheSerGlyPhe 123
 QY 467 TCTTGCTATCTTTTATTGCTCTCGGAAGTTCGATAACTACCGGCAAGAGCGGCT 526
 DB 124 SerThrLeuSerPheIleGlnSerProGlyAspIleLysGluGlnGly----- 139
 QY 527 AGCTGC-----TCTACGGGTAGCTTGTAGTTTGCACAAAATAATGTCAGTTTGTCTTCCAG 580
 DB 140 ---CysLeuTyrSerLysAsnAlaLeuMetLeuLeuAsnAsnTyrValValargPheGlu 158
 QY 581 AAAAAGTTTCAACGGATAATGCGGTGCTATCACCGCAAAACTCTTTTCATTACAGGG 640
 DB 159 GlnAsnGlnSerLysThrLysGlyGlyAlaIleSerGlyAlaAsnValThrIleValGly 178
 QY 641 ACTACAATCTCAGCTCTGTTTCTGAAATACTCTCTCAAGAAAGGCGGAGCCATTACAG 700
 DB 179 AsnTyrAspSerValSerPheTyrGlnAsnAlaAlaThr---PheGlyAlaIleHis 197
 QY 701 ACTTCGATGCCCTTACCATTACTGAAACCAAGGGGAAGTCTCTTTTCTGACATACT 760
 DB 198 SerSerGlyProLeuGlnIleAlaValAsnGlnAlaGluIleArgPheAlaGlnAsnThr 217
 QY 761 TCTTCGGAT---TCTGGAGCTGCAATTTTACAGAAGCCTCGTGACTATTTCTAATAT 817
 DB 218 AlaLysAsnGlySerGlyGlyAlaLeuTyrSerAspGlyAspIleAspIleAspGlnAsn 237
 QY 818 GCTAAAGTTTCTCTTATTGACAATAAGTTCACAGGAGCGAGCTCTCTCAACAACGGGGAT 877
 DB 238 AlaTyrValLeuPheArgGluAsnGlu-----AlaLeuThrThrAlaIle 252
 QY 878 ATGTCAGGAGTCTATCTGCTTATATAAACTAGTACAGATACTAAG----- 925
 DB 253 GlyLysGlyGlyAlaValCysCysLeuProThrSerGlySerSerThrProValProIle 272
 QY 926 GTCACCTCACTGGAATCAGATGTTACTTCTTCAGCAACAATACATACGACACAGCGGA 985
 DB 273 ValThrPheSerAspAsnLysGlnLeuValPheGluArgAsnHisSerIleMetGlyGly 292
 QY 986 GGAGCTATCTATGTGAAAAAGCTCGAACTGGCTTCGGAGGAGCTTACCCCTATTTCAGTAGA 1045
 DB 293 GlyAlaIleTyrAlaArgLysLeuSerIleSerSerGlyGlyProThrLeuPheIleAsn 312
 QY 1046 AATAGTGTCAATGGAGGTACAGCTCTCTAAAGTGGAGCATAGTATCGAAGATAGTGG 1105
 DB 313 AsnIleSerTyrAlaAsnSerGlnAsnLeuGlyGlyAlaIleAlaIleAspThrGlyGly 332
 QY 1106 GAATTGAGTTTATCCCGCATAGTGTGACATGCTTTTATAGGGAATACATCACTTCT 1165
 DB 333 GluIleSerLeuSerAlaGluLysGlyThrIleThrPheGlnGlyAsn-----ArgThr 350
 QY 1166 ACTACTCTTCGGACGACATAGAAGTAGTATCGACTTAGGAACAGAGTCGAAGATGACAGCT 1225
 DB 351 SerLeuProPheLeuAsn-----GlyIleHisLeuLeuGlnAsnAlaLysPheLeuLys 368
 QY 1226 TTGCGTTCTGCTGCTGGTAGGCCATCTACTTCTATGATCCCATCACTACAGATCATCC 1285
 DB 369 LeuGlnAlaArgAsnGlyTyrSerIleGluPheTyrAspProIleThr-----SerGlu 386
 QY 1286 ACACAGTTTACAGATGCTCTTAAAGTAAATGAGATCCCGGACAGTCTCCACTACATAT 1345

```

Db      387 AlaAspGlySerThrGlnLeuAsnIleAsnGlyAspProLysAsnLys-----GluTyr 404
QY      1346 ACAGGAAACATCATCTTCACAGGAGAAAGATTATCAGACAGACAGCGCGAGATTCTAAA 1405
Db      405 ThrGlyThrIleLeuPheSerGlyGluLys-----SerLeuAlaAsnAspProArg 421
QY      1406 AATCTTACTTGAAGCTACTAGCCTGTACTCTTTTCAGAGAGGTACTCTATCTTTTAAAA 1465
Db      422 AspPheLysSerThrIleProGlnAsnValAsnLeuSerAlaGlyTyrLeuValIleLys 441
QY      1466 CATGAGAGTCACTCTCAGACTCAGGCATTCTCAACAGCGAGATTCTCGCTCGAAATG 1525
Db      442 GluGlyAlaGluValThrValSerLysPheThrGlnSerProGlySerHisLeuValLeu 461
QY      1526 GAGGTAGAACTACTCTA---GAACCTGCTGATACATAGCACATAAAACAATTTGTGTCATT 1582
Db      462 AspLeuGlyThrLysLeuIleAlaSerLysGluAspIleAlaIleThrGlyLeuAlaIle 481
QY      1583 AACATCAGTCTTATAGACGGTGCAGAGGCAAGGCAAAA---ATAGAAACCAAGAGTACGTCA 1639
Db      482 AspIleAspSerLeuSerSerSerThrAlaAlaValIleLysAlaAsnThrAlaAsn 501
QY      1640 AAAAATCTGACTTATCTGGAACCATCACTTTATTGGACCGCAGCGGCACGTTTATGAA 1699
Db      502 LysGlnIleSerValThrAspSerIleGluLeuIleSerProThrGlyAsnAlaTyrGlu 521
QY      1700 AATCATAGTTTAAAGAAATCCTCAGTCCTCAGACATCTTTAGAGCTCAAA-----GCTTCT 1753
Db      522 AspLeuArgMetArgAsnSerGlnThrPheProLeuLeuSerLeuGluProGlyAlaGly 541
QY      1754 GCACTGTAAACAAGCAGCAGCA-----GTGACTCCAGATCCTATATGGGT 1798
Db      542 GlySerValThrValThrAlaGlyAspPheLeuProValSerPro----- 556
QY      1799 GAGAAATTCCTATACGGCTATCAGGCACTTGGGGCCCAATTTGTTGG---GGGACAGGG 1855
Db      557 -----HisTyrGlyPheGlnGlyAsnTrp---LysLeuAlaTrpThrGlyThrGly 572
QY      1856 GCTTCTACGACTGCAACCTTCAACTGGAGTAAACTGGCTATATCTCTAATCCCGAGCGT 1915
Db      573 ---AsnLysValGlyGluPhePheTrpAspLysIleAsnTyrLysProArgProGluLys 591
QY      1916 ATCGCTCTTTAGTCCCTAATAGCTTATGGAATGCAATTAAGATATAGATCTCTCTCCAT 1975
Db      592 GluGlyAsnLeuValProAsnIleLeuTrpGlyAsnAlaValAspValArgSerLeuMet 611
QY      1976 TATCTTATGGAGACTGCAACGAGGGTTCAGGAGACCGTGTCTTTTGTGTGTGCTGGA 2035
Db      612 GlnValGlnGluThrHisAlaSerSerLeuGlnThrAspArgGlyLeuTrpIleAspGly 631
QY      2036 TTATCTAACTTCTTCCATPAGGATAGTACAAACACAGACCGGGTTCGCCATTTGAGT 2095
Db      632 IleGlyAsnPhePheHisValSerAlaSerGluAspAsnIleArgTyrArgHisAsnSer 651
QY      2096 GCGGTATGTCTATAGGAGAAACCTACATCTTGTTCAGATAGATCTTCTAGTGTGCA 2155
Db      652 GlyGlyTyrValLeuSerValAsnAsnGluIleThrProLysHisTyrThrSerMetAla 671
QY      2156 TTTTGTGAGCTCTTTGGAAGAGATAGAGACTTCTTGTAGCTTAAGATCAAGGTACAGTC 2215
Db      672 PheSerGlnLeuPheSerArgAspLysAspTyrAlaValSerAsnAsnGluTyrArgMet 691
QY      2216 TACGAGGAACTCTCTATTACAGCAACCAACGAAAC-----TATATC 2257
Db      692 TyrLeuGlySerTyrLeuTyrGlnTyrThrThrSerLeuGlyAsnIlePheArgTyrAla 711
QY      2258 TCTCTCTCTGCAAACTAGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2314
Db      712 SerArgAsnProAsnValAsnValGlyIleLeuSerArgArgPheLeuGlnAsnProLeu 731
QY      2315 GTTCTCTTTTCAGGAACCTTAGCTACACCCATCAGGATACAGTCTGAAACCAAGTAT 2374

```

```

Db      732 MetIlePheHisPheLeuCysAlaTyrGlyHisAlaThrAsnAspMetLysThrAspTyr 751
QY      2375 ACAACATATCTCTACTGTGTAAAGAGAGCTGGGGGAATGATAGTTTCGCTTTTGAATTCGGT 2434
Db      752 AlaAsnPheProMetValLysAsnSerTrpArgAsnAsnCysTrpAlaIleGluCysGly 771
QY      2435 GGAAGAGCTCCGATTTGCTTAGATGAAGTCT---CTATTTGAGCTAGTACATGCCCTTC 2491
Db      772 GlySerMetProLeuLeuValPheGluAsnGlyArgLeuPheGlnGlyAlaIleProPhe 791
QY      2492 ATGAAATTCAGTTCCTATGCATCAGCAAGAGTTCCTTAAAGAACACAGGGAACAGAGCT 2551
Db      792 MetLysLeuGlnLeuValTyrAlaTyrGlnGlyAspPheLysGluThrThrAlaAspGly 811
QY      2552 CGTGAATTTGGAAGTAGCCGCTTTGTGAATCTTCCTATCCCTATCCGGATCCGATTTGAT 2611
Db      812 ArgArgPheSerAsnGlySerLeuThrSerIleSerValProLeuGlyIleArgPheGlu 831
QY      2612 AAGGAATCAGACTCCAGAGTCAACGACGACGATCTACTCTTGTGTTATCTACTGTGATCTT 2671
Db      832 LysLeuAlaLeuSerGlnAspValLeuTyrAspPheSerPheSerTyrIleProAspIle 851
QY      2672 GTTCGTAGTAAACCCGACTGTACGACACACTGCGAATTCAGCGTGTGATCTTGTGAAAC 2731
Db      852 PheArgLysAspProSerCysGluAlaLeuValIleSerGlyAspSerTrpLeuVal 871
QY      2732 TTCGGTACGAATTTGGCAAGACAGCTTTCCTCTCGTCGAGGGAACCATTTTGTCTTT 2791
Db      872 ProAlaAlaHisValSerArgHisAlaPheValGlySerGlyThrGlyArgTyrHisPhe 891
QY      2792 AACTCAATTTTGAAGCTTTAGCCAAATTTCTTTGAATTCGCTGGTGCATCTCGCAAT 2851
Db      892 AsnAspTyrThrGluLeuLeuCysArgGlySerIleGluCysArgProHisAlaArgAsn 911
QY      2852 TACAATGTAGACTTAGGAGCAAAATACCAATTC 2884
Db      912 TyrAsnIleAsnCysGlySerLysPheArgPhe 922

RESULT 9
PM13 CHLPN
ID PM13 CHLPN STANDARD; PRT; 973 AA.
AC Q92856; Q9K2A1; Q9Z4I0,
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, last sequence update)
DT 16-OCT-2001 (Rel. 40, last annotation update)
DE Probable outer membrane protein pmp13 precursor (Polymorphic membrane
protein 13) (Outer membrane protein 14).
GN PMP13 OR OMP14 OR CPN0453 OR CP0299.
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., Baas S.,
Linher K., Hickey E.K., Peterson J., Utterback T., Berry K., Baas S.,
Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406 (2000).
RN [3]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CML029 from USA.";
RL Nucleic Acids Res. 28:2311-2314 (2000).
RN [4]
RP SEQUENCE OF 1-262 FROM N.A.
RC STRAIN=VR1310;
RX MEDLINE=20007584; PubMed=10539856;
RA Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
RA Madсен A.S., Knudsen K., Falk E., Birkelund S.;
RT "Molecular biology of Chlamydia pneumoniae surface proteins and their
RT role in immunopathogenicity.";
RL Am. Heart J. 138:S491-S495 (1999).
CC -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AB001629; RAD18595.1; -.
DR EMBL; AB002191; RAF38156.1; ALT_INIT.
DR EMBL; AP002546; ERA98660.1; -.
DR EMBL; AJ133034; CAB37074.1; -.
DR PIR; B86547; B86547.
DR PIR; F72076; F72076.
DR PHCI-2DPAGE; 09Z896; -.
DR TIGR; CP0299; -.
DR InterPro; IPR005546; Autotransporter.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF02415; DUF145; 3.
DR TIGRPFAMs; TIGR01376; POMP repeat; 7.
DR Outer membrane; Signal; Multigene family; Complete proteome.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 973 PROBABLE OUTER MEMBRANE PROTEIN PMP13.
FT CONFLICT 258 258 N -> Y (IN REF. 4).
SQ SEQUENCE 973 AA; 102761 MW; E02A69F611DBEF2 CRC64;
Alignment Scores:
Pred. No.: 1,72e-75 Length: 973
Score: 1377.50 Matches: 350
Percent Similarity: 49.95% Conservative: 153
Best Local Similarity: 34.76% Mismatches: 391
Query Match: 25.47% Indels: 113
DB: Gaps: 25
US-09-428-122-1 (1-3000) x PMP13_CHLPN (1-973)
QY 101 ATGAAGCTCTTTTCCCAAGTTTGATTCTTACA-----TTTGCTATT 145
Db 1 MetLysThrSerIleArgLysPheLeuIleSerThrThrLeuAlaProCysPheAla--- 19
QY 146 TTCCTTTGCTATGATGCTACCGAGACAGTTTTCGATTCAGTCCGAGTTTCGATGGG 205
Db 20 -----SerThrAlaPheThrValGluValIleMetProSerGluAsnPheAspGly 36
QY 206 AATAAAATGGTAATTTTTCAGTTGCTGAGAGTCAGAGATGCTGGAATCTACCTACCTA 265
Db 37 SerSerGlyLysIlePheProTyThrThrLeuSerAspProArgGlyThrLeuCysIle 56
QY 266 TTTAAGGGAATGCTCACTCTAGAAATATTTCTGGAAACAGCGACCAAGCAATCAAAAAGC 325
Db 57 PheSerGlyAspLeuTyThrIleAlaAsnLeuAspAsnAlaIleSerArgThrSerSer 76

```

```

QY 326 TGTTTTAACAACACTAAGGCGATTTCGACTTTACAGGTAACGGGAACCTCTCTATTGTC 385
Db 77 CysPheSerAsnArgAlaGlyAlaLeuGlnIleLeuGlyLysGlyValPheSerPhe 96
QY 386 CAACACGGTGGATGACGGGACTGTAGCAGGGGCTCTGTAAACAGCAGCGGTGTAGATAAA 445
Db 97 LeuAsnIle---ArgSerSerAlaAspGlyAlaAlaIleSerSerValIleThrGluAsn 115
QY 446 -----TCTACCAGTTTATAGGGTTTTCGCTATCTTTT----- 481
Db 116 ProGluLeuCysProLeuSerPheSerGlyPheSerGlnMetIlePheAspAsnCysGlu 135
QY 482 -----ATTGCGTCTCTGGAGTTTCGATACTACCGCACAAGGAGCGGTAGCTGC 532
Db 136 SerLeuThrSerAspThrSerAlaSerAsnValIleProHisAlaSerAlaIleTyAla 155
QY 533 TCTACGGGTAGCTTGAGTTTGACAAAATGTCAGTTTGCTCTTCAGCAAAAACCTTTTCA 592
Db 156 ThrThrProMetLeu---PheThrAsnAsnAspSerIleLeuPheGlnTyAsnArgSer 174
QY 593 ACGGATATGCGGTGCTATCACCAGAAAACCTCTTCATTAACAGGACTACATGTCA 652
Db 175 AlaGlyPheGlyAlaAlaIleArgGlyThrSerIleThrIleGluAsnThrLysLysSer 194
QY 653 GCTCTGTTTCTGAAATACCTCTCAAGAAAGCGGAGCATTACAGACTTCGATGCC 712
Db 195 LeuLeuPheAsnGlyAsnGlySerIleSerAsnGlyAlaLeuThrGlySerAlaAla 214
QY 713 CTTACCACTTACTGAAACCAAGGGGAAGTCTCTTTTCTGAC----- 754
Db 215 IleAsnLeuIleAsnAsnSerAlaProValIlePheSerThrAsnAlaThrGlyIleTy 234
QY 754 ----- 754
Db 235 GlyGlyAlaIleTyLeuThrGlyGlySerMetLeuThrSerGlyAsnLeuSerGlyVal 254
QY 755 -----AATACTTCTTCGGATTCTCGAGCTGCAATTTTACAGAAGCCTCGGTGACT 805
Db 255 LeuPheValAsnAsnSerSerArgSerGlyAlaIleTyAlaAsnGlyAsnValThr 274
QY 806 ATTTCTAATAATGCTTAAAGTTTCTCTTTATGACATAAGGTCACAGAGCAGAGTCC--- 862
Db 275 PheSerAsnAsnSerAspLeuThrPheGlnAsnAsnThrAlaSerProGlnAsnSerLeu 294
QY 863 -----TCAACAACGGGGATATGTCA---GGAGGT 889
Db 295 ProAlaProThrProProThrProProAlaValThrProLeuLeuGlyTyGly 314
QY 890 GCTATC---TGT-----GCTTATAAACTAGTACAGATACTAAGGTCAACCTCACT 937
Db 315 AlaIlePheCysThrProProAlaThrProProThrGlyValSerLeuThrIleSer 334
QY 938 GGAAATCAGATGTTACTCTTCAGCAACAATACATCGACAACAGCGGAGGAGCTATCTAT 997
Db 335 GlyGluAsnSerValThrPheLeuGluAsnIleAlaSerGluGlnGlyGlyAlaLeuTy 354
QY 998 GTGAAAAGCTCGAAGTCTCCGAGAGCTTACCTATTCAGTAGAAAATAGTCTCAAT 1057
Db 355 GlyLysLysIleSerIleAspSerAsnLysSerThrIlePheLeu----- 369
QY 1058 GGAGGTACAGCTCTTAAAGTGGAGCGCATAGTATCGAAGATAGTGGGAATTTAGATTTA 1117
Db 370 GlyAsnThrAlaGlyGlyGlyAlaIleAlaIleProGluSerGlyGluLeuSerLeu 389
QY 1118 TCCGCCGATAGTGGTACATTGCTTTTTTAGGGAAT---ACAGTCACCTTCTACTCTCT 1174
Db 390 SerAlaAsnGlnGlyAspIleLeuPheAsnLysAsnLeuSerIleThrSerGlyThrPro 409
QY 1175 GGGACGAATAGAAAGTATGATTCAGCTTAGAAGAGTGCAGAAATGACAGCTTTGGCTCT 1234
Db 410 -----ThrArgAsnSerIleHisPheGlyLysAspAlaLysPheAlaThrLeuGlyAla 427
QY 1235 GCTGCTGGTAGGCCATCTACTTCTATGATCCATACTACAGGA-----TCATCCACA 1288

```

```

Db      428 ThrGlnGlyThrLeuThrPheThrAspProIleThrSerAspLeuSerAlaala 447
      ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1289 ACAGTTACAGAGTCTTAAAGTTAAT---GAGACTCCGGCAGATTCGCACTACAAAT 1345
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      448 SerAlaAlaThrValValAlaSerProLysAlaSerAlaAspGlyAla-----Tyr 465
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1346 ACAGGGACATCATCTTCACAGGAGAAAGTTATCAGACAGACAGCCGACATTCCTAAA 1405
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      466 SerGlyThrIleValPheSerGlyGluThrLeuThrAlaThrGluAlaAlaThrProAla 485
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1406 AATCTTACTTCAAGCTACTACAGCTGTAACTCTTTTCAGGAGGTACTCTATCTTTAAAA 1465
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      486 AsnAlaThrSerThrLeuAsnGlnLysLeuGluLeuGluGlyThrLeuAlaLeuArg 505
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1466 CATGAGTGAAGTCTGAGACTCAGGCATTCATCAACAGGCAGATTCCTGCTCGAAATG 1525
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      506 AsnGlyAlaThrLeuAsnValHisAsnPheThrGlnAspGluLysSerValValIleMet 525
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1526 GACGTAGGAACACTCTAGAACCTGCTGATACG-----ACC 1564
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      526 AspAlaGlyThrThrLeuAlaThrThrAsnGlyAlaAsnAsnThrAspGlyAlaIleThr 545
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1565 ATAAACAATTTGGTCAATTAACATCATGTTCTATAGACGGTGCAGGAGCAAAATAGAA 1624
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      546 LeuAsnLysLeuValIleAsnLeuAspSerLeuAspGlyThrLysAlaAlaValValAsn 565
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1625 ACCAAGACTACGTCAAAAATCTGATTTATCGAACCATCATCTTATTGACCCGACG 1684
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      566 ValGlnSerThrAsnGlyAlaLeuThrIleSerGlyThrLeuGlyLeuValLysAsnSer 585
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1685 GGCACGTTTATGAAATCATAGTTTAAGAAATCCT-----CAGTCCTACGACATCTTA 1738
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      586 GlnAspCysCysAspAsnHisGlyMetPheAsnLysAspLeuGlnGlnValProIleLeu 605
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1739 GAGCTCAAGCT---TCTGGAACTGTAAACAAGCAGCGAGTCAGCTCCAGATCCTATAATG 1795
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      606 GluLeuLysAlaThrSerAsnThrValThrThrAspPheSerLeuGlyThrAsnGly 625
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1796 GGTGAGAAATCCATTACGGCTATCAGGGAATCTGGGGCCCAATTTGTTGGGGACAGG 1855
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      626 TyrGlnGlnSerProGlyThrGlyThrGlnGlyThrTrpGluPheThrIleAspThrThr--- 644
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1856 GCTTCTACGACGCAACCTTCACTGCTAACTGCTATATTCCTAATCCCGAGCT 1915
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      645 ---ThrHisThrValThrGlyAsnTrpLysLysThrGlyThrLeuProHisProGluArg 663
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1916 APCGGCTCTTTAGTCCCTAATAGCTTATGGAATGCAATTTATAGATATTAGCTCTCCAT 1975
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      664 LeuAlaProLeuIleProAsnSerLeuTrpAlaAsnValIleAspLeuArgAlaValSer 683
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1976 TATCTTATGGAGACTGCAAACGAGGTGTCAGGAGACCGCTGCTTTTGGTGGCTGGA 2035
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      684 GlnAlaSerAlaAlaAspGlyGluAspValProGly---LysGlnLeuSerIleThrGly 702
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2036 TTATCTAATCTTCCATAAGATAGTAGTACAAAACACGACGCGGGTTTCGCCATTTGACT 2095
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      703 IleThrAsnPhePheHisAlaAsnHisThrGlyAspAlaArgSerThrArgHisMetGly 722
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2096 GCGGCTTATGTCATAGGAGGAAACCTACATCTTGTTCAGATAAGATCTTATAGTGTGCA 2155
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      723 GlyGlyThrLeuIleAsnThrThrArgIleThrProAspAlaAlaLeuSerLeuGly 742
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2156 TTTTGTACGCTTTTGGAGAGATAGACATCTCTTTGTAGCTTAAGATCAAGGTACAGTC 2215
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      743 PheGlyGlnLeuPheThrLysSerLysAspThrLeuValGlyHisGlyHisSerAsnVal 762
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2216 TACGGAGGAACTCTATTACAGCACACGAAACCTATATCTCTCTCTCTCTCTCTCTCTCT 2275
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      763 TyrPheAlaThrValThrSerAsnIleThrLysSerLeuPheGly----- 777
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2276 CGGCGCTTGTCTGCTTATGCTTCTACAGAGATTCCTGTTCTCTTTTCAGGAAACCTT 2335
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db      778 -----SerSerArgPhePheSerGlyThr 786
QY      2336 AGC-----TACACCATACGGATTAACATCTGAAACCAAGATATACAAATATCTCT 2386
      ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      787 SerArgValThrThrSerArgSerAsnGluLysValLysThrSerThrLysLeuPro 806
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2387 ACTGTTAAAGGAAGCTGGGGAATCATAGTTTCGCTTTTAGATTCGGTGGGAGAGCTCG 2446
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      807 LysGlyArgCysSerTrpSerAsnAsnCysTrpLeuGlyGluLeuGluGlyAsnLeuPro 826
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2447 ATTTCTTATAGATGAAAGTGCTCTA---TTTGACAGATACATGCCCTTCATGAAATTCGAG 2503
      ||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      827 IleThrLeuSerSerArgIleLeuAsnLeuLysGlnIleProPheValLysAlaGlu 846
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2504 TTTGCTATGACATCAGAAAGGTTTAAAGAACAGGGAACAGAGCTCGTGAATTTGGA 2563
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      847 ValAlaThrAlaThrHisGlyGlyIleGlnGluAsnThrProGluGlyArgIlePheGly 866
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2564 AGTAGCCGCTTGTGAATCTTCCCTTACCTATCGGATCCGATTTGATAGGAATCAGAC 2623
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      867 HisGlyHisLeuLeuAsnValAlaValProValGlyValArgPheGlyLysSerHis 886
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2624 TGCCAAGATGCAACGTCACAATCTTACTCTTGGTTTACTGTGATCTTGTCTAGTAAC 2683
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      887 AsnArgProAspPheThrThrIleValAlaThrAlaProAspValThrArgHisAsn 906
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2684 CCGGCTCTACGACAACTCGGATAGCGGTGATCTTGGAAACCTTCGGTACGAAT 2743
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      907 ProAspCysAspThrThrLeuProIleAsnGlyAlaThrTrpThrSerIleGlyAsnAsn 926
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2744 TTGGCAACACAAAGCTTTAGTCTCTCGTCGCGGACCACTTTTGTCTTAACTCAATTTT 2803
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      927 LeuThrArgSerThrLeuLeuValGlnAlaSerSerHisThrSerValAsnAspValLeu 946
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2804 GAAGCTTTAGCCAAATTTCTTTGAATTCGGTGGGTCTATCTCGCAATTAACAATCTAGAC 2863
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      947 GluIlePheGlyHisCysGlyCysAspIleArgThrSerArgGluThrThrLeuAsp 966
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2864 TTAGAGCAAAATACCAATTC 2884
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      967 IleGlySerLysLeuArgPhe 973

RESULT 10
PMPG_CHLTR
ID PMPG_CHLTR STANDARD; PRT; 1013 AA.
AC 084879;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Probable outer membrane protein pmpG precursor (Polymorphic membrane
DE protein G).
GN PMPG OR C7871.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
RX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/Cx;
RC MEDLINE=9900809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis.";
RL Science 282:754-759(1998).
CC -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AB001360; AAC68469.1; -
 DR PIR; G71460; G71460;
 DR PHCI-2DPAGE; O84879; -
 DR InterPro; IPR006315; Autotransport.
 DR InterPro; IPR005546; Autotransporter.
 DR InterPro; IPR003368; Chlamydia PMP.
 DR Pfam; PF03797; Autotransporter; 1.
 DR Pfam; PF02415; DUPL45; 1.
 DR TIGRfams; TIGR01414; autotrans_bar1; 1.
 DR TIGRfams; TIGR01376; POMP_repeat; 6.
 KW Outer membrane; Signal; Multigene family; Complete proteome.
 FT SIGNAL 1 27 POTENTIAL.
 FT CHAIN 28 1013 PROBABLE OUTER MEMBRANE PROTEIN PMPG.
 SQ SEQUENCE 1013 AA; 107366 MW; F0927743C0A651DD CRC64;

Alignment Scores:

Pred. No.: 4,77e-60 Length: 1013
 Score: 1120.50 Matches: 313
 Percent Similarity: 44.66% Conservative: 160
 Best Local Similarity: 29.56% Mismatches: 409
 Query Match: 20.72% Indels: 177
 DB: 1 Gaps: 33

US-09-428-122-1 (1-3000) x PMPG_CHLTR (1-1013)

QY 101 ATGAAGTCTCTTTCCCAAGTTGTATTTCTACATTTGCTATTTTCCTCTGTCTATG 160
 DB 1 MetGlnThrSerPheHisLysPhePheLeuSerMetIleuAlaTyrSerCysSer 20
 QY 161 ATTGCT-----ACCGACACAGTTTGGATTCAAGTCGAGTTTCGATGGGAAT 208
 DB 21 LeuSerGlyGlyTyrAlaAlaGluIleMetIleProGlnGlyIleTyrAspGlyGlu 40
 QY 209 AAA---AATGGTAATTTTTCAGTTCGTGAGAGTCAGGAAGATCGTGAACCTACCTA 265
 DB 41 ThrLeuThrValSerPheProTyrThrValIleGlyAspProSerGlyThrThrValPhe 60
 QY 266 TTTAAGGGAAATGTCACCTAGAAATATTCCTGGAAACGACGACAGCAATCACAAAAGC 325
 DB 61 SerAlaGlyGluLeuThrLeuLysAsnLeuAspAsnSerIleAlaLeuProLeuSer 80
 QY 326 TGTTTTAACACACTAAGGGCGATTGACTTTACAGTAACGGGAACCTCTATTGTTTC 385
 DB 81 CysPheGlyAsnLeuLeuGlySerPheThrValLeuGlyArgGlyHisSerLeuThrPhe 100
 QY 386 CAACAGGTGATCAGGAGCTGTAGCAGGGCTGCTGTTAACAGCGGTGGTAGATAAA 445
 DB 101 GluAsnIleArgThrSerThr---AsnGlyAlaAlaLeuSerAspSerAlaAsnSerGly 119
 QY 446 TCTACCAAGCTTATAGGGTTTCTTCGCTATCTTT----- 481
 DB 120 LeuPheThrIleGluGlyPheLysGluLeuSerPheSerAsnCysAsnSerLeuLeuAla 139
 QY 482 -----ATTCCGCTCTCGGAAGTTTCGATAACTPACC 511
 DB 140 ValLeuProAlaAlaThrThrAsnAsnGlySerGlnThrProThrThrSerThrPro 159
 QY 512 GGCAAGGCGGTAGTCTCTACGGTAGCTTGGATTGACAAAATGTCAGTTG 571
 DB 160 SerAsnGlyThrIleTyrSerLysThr---AspLeuLeuLeuLeuAsnGluLysPhe 178
 QY 572 CTCCTTCAGAAAACCTTTCAACCGGATAATGGCGGTCTATCACCGCAAAACCTTTCA 631
 DB 179 SerPheTyrSerAsnLeuValSerGlyAspGlyAlaIleAspAlaLysSerLeuThr 198
 QY 632 TTAACAGGACTACAATGTACGCTCTGTTTCTGAAATACCTCTCAAGAAAGCGCGA 691
 DB 199 ValGlnGlyIleSerLysLysCysValPheGlnGluAsnThrAlaGlnAlaAspGlyGly 218

QY 692 GCCATTCCAGACTTCGATCCGCTTACCATTACTGGAACCAAGGGAAGTCTCTTT--- 748
 DB 219 AlaCysGlnValValThrSerPheSerAlaMetAlaAsnGluAlaProIleAlaPheIle 238
 QY 748 ----- 748
 DB 239 AlaAsnValAlaGlyValArgGlyGlyIleAlaAlaValGlnAspGlyGlnGlnGly 258
 QY 749 -----TCTGACATPACTTCTTCGATTCT--- 772
 DB 259 ValSerSerSerThrThrGluAspProValValSerPheSerArgAsnThrAlaVal 278
 QY 773 -----GGAGCTGCAATTTTACAGAGCCTCGCTGACT 805
 DB 279 GluPheAspGlyAsnValAlaArgValGlyGlyIleTyrSerTyrGlyAsnValAla 298
 QY 806 ATTTCTTAATAAGTCTAAAGTTTCTTTTATGTACAAT----- 841
 DB 299 PheLeuAsnAsnGlyLysThrLeuPheLeuAsnValAlaSerProValTyrIleAla 318
 QY 842 -----AGGTCCAGAGCGAGCTCTCAACACAGGGGGATATG-----TCAGAGGT 889
 DB 319 AlaGluGlnProThrAsnGlyGlnAlaSerAsnThrSerAspAsnTyrGlyAspGly 338
 QY 890 GCTATC---TGT-----GCTTATAAACTAGTACAGATTAAGGTCAAC 931
 DB 339 AlaIlePheCysLysAsnGlyAlaGlnAlaAlaGlySerAsnAsnSerGlySerValSer 358
 QY 932 CTCACCTGGAATCAGATGTTCTCTCAACAATATACATCGACACAGCGGAGAGT 991
 DB 359 PheAspGlyGluGlyValValPhePheSerSerAsnValAlaAlaGlyLysGlyAla 378
 QY 992 ATCTATGTGAAAAGCTCGAAGTCTCCGAGGACTTACCTATTCTAGTAAATAGT 1051
 DB 379 IleTyrAlaLysLysLeuSerValAlaAsnCysGlyProValGlnPheLeuGlyAsnIle 398
 QY 1052 GTCAATGGAGGTACAGCTCTCTAAAGTGGAGCATAGTATCGAAGATAGTGGGAATG 1111
 DB 399 AlaAsn-----AspGlyGlyAlaIleTyrLeuGlyGluSerGlyGluLeu 413
 QY 1112 AGTTTATCCGCCGATAGTGTGACATTTCTTTTAGGGAAT-----ACAGTCACT 1162
 DB 414 SerLeuSerAlaAspTyrGlyAspIleIlePheAspGlyAsnLeuLysArgThrAlaLys 433
 QY 1163 TCTACTACTCTCTGGGAGGAAT-----AGAAGTAGTATCGACTTAGGAACG 1207
 DB 434 GluAsnAlaAlaAspValAsnGlyValThrValSerSerGlnAlaIleSerMetGlySer 453
 QY 1208 AGTGCAAGATGACAGCTTTGCTGCTGCTGGTAGAGCCATCTACTTATGATCCC 1267
 DB 454 GlyGlyIleIleThrThrLeuArgAlaLysAlaGlyHisGlnIleLeuPheAsnAspPro 473
 QY 1268 ATAACCTACAGGATTCCTCCACA-----GTTACAGATGCTTTAAAGTTAAT 1315
 DB 474 IleGluMetAlaAsnGlyAsnAsnGlnProAlaGlnSerSerGluProLeuLysIleAsn 493
 QY 1316 GAGACTCCGCGAGATTCTGCCTACATATACAGGGAACATCATCTCTCAGAGAAAAG 1375
 DB 494 AspGlyGluGly-----TyrThrGlyAspIleValPheAla----- 505
 QY 1376 TTATCAGACAGACAGCGCGAGATTCTAAATAATCTTCTCGAAGCTACTACAGCCTGTA 1435
 DB 506 -----AsnGlyAsnSerThrLeuTyrGlnAsnVal 515
 QY 1436 ACTCTTTTCAGGAGTACTCTATCTTTTAAACATGGAGTACTCTCGAGACTCAGGATTC 1495
 DB 516 ThrIleGluGlnGlyArgIleValLeuArgGluLysAlaLysLeuSerValAsnSerLeu 535
 QY 1496 ACTCAACAGCAGGATCTCTGCTCGAAATGGAGCTAGGAACTACTCTAGAA----- 1546
 DB 536 SerGlnThrGlyGlySer---LeuTyrMetGluAlaGlySerThrLeuAspPheValThr 554
 QY 1547 -----CCTGCTGATACTAGC-----ACCAATAACAATTTGGTC 1579

```

Db      555 ProGlnProGlnGlnProProAlaAlaAsnGlnLeuLeuThrLeuSerAsnLeuHis 574
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1580 ATTAACATCAGTTCTATA-----GACGGTGCA-----1606
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      575 LeuSerLeuSerLeuLeuAlaAsnAsnAlaValThrAsnProProThrAsnProPro 594
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1607 -----AAGAGCGCAAAATAGAACCAAGCTAGCTCAAAAATCTGACTTTATCT 1657
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      595 AlaGlnAspSerHisProAlaIleIleGlySerThrThrAlaGlySerValThrIleSer 614
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1658 GGAACCATCAGTTTATGACCGCGGCGGACGGTCTTTATGAAAATCATAGT---TTAAGA 1714
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      615 GlyProIlePhePheGluAspLeuAspAspThrAlaIleAspArgGlyAspTrpLeuGly 634
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1715 AATCTCTCAGCTCAGACATCTTAGAGCTCAAGCTCTCGAACTGTAACAGACCCGCA 1774
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      635 SerAsnGlnLysIleAspValLeuLysLeuGlnLeu---GlyThrGlnProSerAlaAsn 653
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1775 GTGACTCCAGATCCCTATAATAGGCGTGAAGAAATTC---CATTACGGCTATCAGGAACTTGG 1831
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      654 AlaProSerAspLeuThrLeuGlyAsnGluMetProLysTyrGlyTyrGlnGlySerTrp 673
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1832 GCGCCCAATTGTTGGGGACAGGGCTCTACAGCTGCAACCTTC-----AAC 1879
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      674 ---LysLeuAlaIleAspProAsnThrAlaAsnAsnGlyProIleThrLeuLysAlaThr 692
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1880 TGGACTAAACCTGCTATATCTCTATCCCGAGCGCTATCGGCTCTTTAGTCCCTAATAGC 1939
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      693 TrpThrLysThrGlyTyrAsnProGlyProGluArgValAlaSerLeuValProAsnSer 712
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      1940 TTATGAATGCAATTTATAGATTTAGCTCTCTCCATTATCTTATGGAGACTGCAACGAA 1999
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      713 LeuTrpGlySerIleLeuAspIleArgSerAlaHisSerAlaIleGlnAlaSerValAsp 732
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2000 GGGTTCAGGGAGACGGCTCTTTGGTGTGCTGATTTATCTTAACTCTTCCATAAGGAT 2059
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      733 GlyArgSerTyrCysArgGlyLeuTrpValSerGlyValSerAsnPheThrIleHisAsp 752
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2060 AGTCAAAAACACGACGCGGTTTCGCCATTTGAGTGGCGGTTATGTCATAGAGAGAAC 2119
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      753 ArgAspAlaLeuGlyGlnGlyTyrArgTyrIleSerGlyGlyTyrSerLeuGlyAlaAsn 772
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2120 CTACATCTGTTTCAGATTAAGATCTTAGTGTGCTGCAATTTGTCAGCTCTTTGAGAGAT 2179
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      773 SerTyrPheGlySer---SerMetPheGlyLeuAlaPheThrGluValPheGlyArgSer 791
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2180 AGAGACTACTTTGCTAGCTAAGATCAAGGTACAGTCTACGGAGGAACCTCTATTACCAG 2239
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      792 LysAspTyrValValCysArgSerAsnHisAlaCysIleGlySerValTyr----- 809
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2240 CACAAGAAACCTATATCTCTCTTCCTGCAAACTACGGCTTGTGTTGCTTATGTT 2299
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      810 -----LeuSerThrLysGlnAlaLeuLysCysGly---SerTyr--- 820
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2300 CCTACAGAGATTCCTGTTCTCTTTTCAGGA-----AACCCTAGCTACCCCATAGC 2350
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      821 -----LeuPheGlyAspAlaPheIleArgAlaSerTyrGlyPheGly 834
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2351 GATAACGATCTGAAAACCAAGATATACACATATCTCTGTTAAAGGAAGCTGGGGGAAT 2410
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      835 AsnGlnHisMetLysThrSerTyrThrPheAlaGluSerAspValArgTrpAspAsn 854
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2411 GATAGTTTCGCTTTAGAAATTCGGTGGAGAGCTCCGATTGCTTAGATGAAGTGTCTTA 2470
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      855 AsnCysLeuValGlyGluIleGlyValGlyLeuProIleValIleThrProSerLysLeu 874
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2471 TTT---GAGCAGTACATGCGCTTCATGAAATTCAGTTGCTCTATGCAATCAGGAGGT 2527
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      875 TyrLeuAsnGluLeuArgProPheValGlnAlaGluPheSerTyrAlaAspHisGluSer 894
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      2528 TTTAAGACACAGGACAGAGCTGCTGATTTGGAGTAGCGCTTGTGATCTTGGC 2587
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 11

```

PMPG CHLMU
ID_PMPG CHLMU STANDARD; PRT; 987 AA.
AC Q9PJ45;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable outer membrane protein pmpG precursor (Polymorphic membrane
DE protein G).
GN PMPG OR TC0263.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]_SEQUENCE FROM N.A.
RP STRAIN=MoPn / Nig9;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RA "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE002293; AAF39132.1; -.
CC PIR; H81722; H81722.
CC TIGR; TC0263;
CC InterPro; IPR006315; Autotransport.
CC InterPro; IPR005546; Autotransporter.
CC InterPro; IPR003368; Chlamydia PMP.
CC Pfam; PF03797; Autotransporter; 1.
CC Pfam; PF02415; DUF145; 1.
CC TIGRFAMs; TIGR01414; autotrans_bar1; 1.
CC TIGRFAMs; TIGR01376; POMP_repeat; 6.
CC Outer membrane; Signal; Multigene family; Complete proteome.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 987 PROBABLE OUTER MEMBRANE PROTEIN PMPG.
FT SEQUENCE 987 AA; 104867 MW; 32079BD6BEB2DA42 CRC64;
SQ

```

Alignment Scores: 7,11e-56 Length: 987
Pred. No.: 1051,00 Matches: 294
Score: 43.20% Conservative: 154
Percent Similarity: 28.35% Mismatches: 429
Best Local Similarity: 19.43% Indels: 160
Query Match: 1 Gaps: 26
DB: 1

US-09-428-122-1 (1-3000) x PMPG_CHLMU (1-987)

QY 101 ATGAAGTCTCTTCCCAAGTTGTATTTCTACATTTGCTATTTTCCCTTGTCTANG 160
DB 2 MetGlnThrProPheHisLysPhe-----PheLeuLeuAlaMet 14
QY 161 ATTGCTACCGAGACAGTTTGGATTCAAGTCCAGT----- 196
DB 15 LeuSerTySerLeuLeuGlnGlyHisAlaAlaAspIleSerMetProGlyIle 34
QY 197 TTCGATGGGAATAAA---AATGGTAATTTTTCAGTTTCGTGAGAGTCAGGAAGATGCTGA 253
DB 35 TyrAspGlyThrThrLeuThrAlaProPheProTyThrValIleGlyAspProArgGly 54
QY 254 ACTACCTACCTATTAAAGGAATATGCTCTAGAAAATATTCCTGGAACAGCCAGCA 313
DB 55 ThrLysValThrSerSerGlySerLeuLeuLeuLysAsnLeuAspAsnSerIleAlaThr 74
QY 314 ATCACAAAAAGCTGTTTAAACAACACTAAGGGCGATTGACTTTTCAGGTAACGGGAAC 373
DB 75 LeuProLeuSerCysPheGlyAsnLeuLeuGlyAsnPheThrIleAlaGlyArgGlyHis 94
QY 374 TCTCTATTGTTCCAAACGGGATGTCAGGAGCTGTAGACGGGCTGCTGTTAAACAGCAGC 433
DB 95 SerLeuValPheGluAsnIleArgThrSerThr---AsnGlyAlaAlaLeuSerAsnHis 113
QY 434 GTGGTAGATAAACTACCAGCTTTATAGGTTTCTTCGTATCTTTATT----- 484
DB 114 AlaProSerGlyLeuPheValIleGluAlaPheAspGluLeuSerLeuLeuAsnCysAsn 133
QY 485 -----GGCTCTCTCGGAAGTTTCGATACTACCGGCAAAAGGAGCC 523
DB 134 SerLeuValSerValValProGlnThrGlyGlyThrThrSerValProSerAsnGly 153
QY 524 GTTAGCTGCTCTACGGGTAGCTTTGACAAAAAATGCTATGCTTTGCTCTTCAGCAAA 583
DB 154 ThrIleTySerArgThrAspLeuValLeuArgAspIleLysLysValSerPheTySer 173
QY 584 AACTTTTCAACGGATAATGGGGTGTCTATCACCGCAAAACTCTTTTCATTAAACAGGACT 643
DB 174 AsnLeuValSerGlyAspGlyGlyAlaIleAspAlaGlnSerLeuMetValAsnGlyIle 193
QY 644 ACAATGTCAGCTCTGTTTCTGAAAATACCTCTCTCAAGAAAGCGGAGCCATTTCAGACT 703
DB 194 GluLysLeuCysThrPheGlnGluAsnValAlaGlnSerAspGlyGlyAlaCysGlnVal 213
QY 704 TCCGATGCCCTTACCATCTACTGGAACCAAGGGAAGTCTCTTTTCTGACATCTTCT 763
DB 214 ThrLysThrPheSerAlaValAlaGlyAsnLysValProLeuSerPheLeuGlyAsnValAla 233
QY 764 TCGGATTCTGGAGCTGCNAAT-----TTTACA 790
DB 234 GlyAsnLysGlyGlyGlyValAlaAlaValLysAspGlyGlnGlyAlaGlyGlyAlaThr 253
QY 791 GAAGCTCGGGTACTATTCTTAATAATGCTAAAGTTTCTTTTATGACAATTAAGTCT--- 847
DB 254 AspLeuSerValAsnPheAlaAsnAsnThrAlaValGluPheGluGlyAsnSerAlaArg 273
QY 848 ACAGGAGCGAGCTCTCTCAACACCGGGGATATGTC----- 883
DB 274 IleGlyGlyGlyIleTySerAspGlyAsnIleSerPheLeuGlyAsnAlaLysThrVal 293
QY 883 ----- 883

DB 294 PheLeuSerAsnValAlaSerProIleTyValAspProAlaAlaAlaGlyGlyGlnPro 313
QY 884 -----GGAGTGTCTATCTGTGCTTATAAACTAGTACA 916
DB 314 ProAlaAspLysAspAsnTyArgGlyGlyAlaIlePheCysLysAsnAspThrAsn 333
QY 917 GATACTAAGTCAACCTCAGTGAATCAGATGTTACTTTCAGCAACAATACATCGACA 976
DB 334 IleGlyGluValSerPheLysAspGluGlyValValPhePheSerLysAsnIleAlaAla 353
QY 977 ACAGCGGAGGAGCTATCTATGTGAAAAGCTCGAACTGGCTTCGAGGAGCTTACCCCTA 1036
DB 354 GlyLysGlyGlyAlaIleTyAlaLysLysLeuThrIleSerAspCysGlyProValGln 373
QY 1037 TTCAGTAGAATAGTGTCAATGAGGTACAGTCCCTAAAGTGGAGCCATAGCTATCGAA 1096
DB 374 PheLeu-----GlyAsnValAlaAsnAspGlyGlyAlaIleTyLeuVal 388
QY 1097 GATAGTGGGAATTGAGTTATCCCGCATAGTGTGACATGCTTTTAGGGAATACA 1156
DB 389 AspGlnGlyGluLeuSerLeuSerAlaAspArgGlyAspIlePheAspGlyAsnLeu 408
QY 1157 GTCACCTTCTACTACTCCTGGG-----ACGAATAGAACTAGT 1192
DB 409 LysArgMetAlaThrGlnGlyAlaAlaThrValHisAspValMetValAlaSerAsnAla 428
QY 1193 ATCGACTTAGGAACGAGTGCAAAGATGACAGCTTTGCTGCTGCTGCTGAGCCATC 1252
DB 429 IleSerMetAlaThrGlyGlyGlnIleThrThrLeuArgAlaLysGluGlyArgGly 448
QY 1253 TACTCTATGATCCCATACTACAGATCATCCACACAGTTACAGATGCTCTTAAAGTT 1312
DB 449 LeuPheAsnAspProIleGluMetAlaAsnGlyGlnProValIleGlnThrLeuThrVal 468
QY 1313 AATGAGACTCCGCGAGATTCTGCACTACAAATATACAGGGAACATCATCTTCACAGAGAA 1372
DB 469 AsnGluGlyGluGly-----TyThrGlyAspIleValPhe----- 480
QY 1373 AAGTTATCAGACAGAGCGCCGAGATCTCTAAATCTTACTTCGAAGCTACTACAGCT 1432
DB 481 -----AlaLysGlyAspAsnVal-----LeuTySerSer 490
QY 1433 GTAACCTTTTCAGGAGTACTCTATCTTTAAACATGGAGTGACTCTCGCAGACTCAGGCA 1492
DB 491 IleGluLeuSerGlnGlyArgIleIleLeuArgGluGlnThrLysLeuValAsnSer 510
QY 1493 TTCACCTCAACAGCAGCATCTCTGCTCGAAATGGACGTAGGAACTACTCTAGAA----- 1546
DB 511 LeuThrGlnThrGlyGlySer---ValHisMetGluGlyGlySerThrLeuAspPheAla 529
QY 1547 -----CCTGCTGATCTAGCACCATTAACAAATTTGTCATTAACATCAGT 1591
DB 530 ValThrThrProAlaAlaAsnSerMetAlaLeuThrAsnValHisPheSerLeuAla 549
QY 1592 TCTATA-----CACGTGCAAAAGAG-----GCAAAATAGAA 1624
DB 550 SerLeuLeuLysAsnAsnGlyValThrAsnProProThrAsnProProValGlnValSer 569
QY 1625 ACCAAAGCTACGTCAAAATAAT-----CTGACTTTTCTGGAACCATCAGT 1669
DB 570 SerProAlaValIleGlyAsnThrAlaAlaGlyThrValThrIleSerGlyProIlePhe 589
QY 1670 TTATGGACCCGACGGCGACGTTTATGAAATCATAGT---TTAGAATCTCTAGTCC 1726
DB 590 PheGluAspLeuAspGluThrAlaTyAspAsnAsnGlnTrpLeuGlyAlaAspGlnThr 609
QY 1727 TAGCATCTTACAGCTCAAGCTTCTGGAACTGTAACAGCACCCAGTGCCTCCAGAT 1786
DB 610 IleAspValLeuGlnLeuHisLeuGlyAlaAsnProProAlaAsnAlaProThrAspLeu 629
QY 1787 CCTATATGGGTGAGAAATTCCTATACGGCTATCAGGAACTTGGGCGCCCAATTTGTTGG 1846
DB 630 ThrLeuGlyAsnGluSerSerLysTyGlyGlnGlySerTrp---ThrLeuGlnTrp 648


```
RESULT 12
PMPI_CHLMU
ID PMPI_CHLMU STANDARD; PRT; 867 AA.
AC Q9PLA1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable outer membrane protein pmpl precursor (Polymorphic membrane
protein 1).
GN PMPI OR TC0267.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MoPn / Nigg;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heideberg J.F.,
White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
Eisen J., Fraser C.M.;
RA "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
pneumoniae AR39."
RT Nucleic Acids Res. 28:1397-1406(2000).
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -|- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC -|- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
DR EMBL; AE002294; AAF39136.1; -.
DR F81721; F81721.
DR TIGR; TC0267; -.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF02415; DUF145; 1.
DR TIGRFAMs; TIGR01376; POMP repeat; 6.
KW Outer membrane; Signal; Multigene family; Complete proteome.
FT SIGNAL 1 29
FT CHAIN 30 867 PROBABLE OUTER MEMBRANE PROTEIN PMPI.
FT SEQUENCE 867 AA; 95017 MW; 557994185A9B5652 CRC64;
Alignment Scores:
Pred. No.: 1.79e-47 Length: 867
Score: 911.00 Matches: 281
Percent Similarity: 44.60% Conservative: 144
Best Local Similarity: 29.49% Mismatches: 384
Query Match: 16.84% Indels: 144
DB: 1 Gaps: 30
US-09-428-122-1 (1-3000) x PMPI_CHLMU (1-867)
QY 149 CTTTGTCTATGATTCACGAGACAGTTCGATTCAGTCGAGTTTCGATGGGAAT 208
Db 18 ProThrAlaIleLeuPheGlyGlnAspAlaLeuAspLysSerAlaLeuIleThrLysAsn 37
QY 209 AAAATGGT-----AATTTTCA----- 226
Db 38 ProAsnSerIleValCysThrPheLeuGluAspCysThrMetGluAsnPheSerProAla 57
QY 227 ---GTTCTGTGAGTCAGCAAGAT-----GCTGGACTACCTACTATT 268
Db 58 LeuLeuSerHisAlaArgGlnAspProLeuTyrllelleGlyAsnThrHis----- 75
QY 269 AAGGAAATGTCACTAGAAAATATT---CCTGGAACAGGCACAGCAATCACAAGAGC 325
Db 76 -----AsnTrpPheValSerAsnLeuHisProSerThr-----AsnGluGlu 89
```

QY 326 TGTGTTTAAACAACACTAAGGGGAGTTGACTTTTACAGGTAACGGGAAGTCTCTATTGTC 385
Db |||||
QY 90 ArgPheLeuLysGluLysGlyAspLeuSerIle----- 100
Db |||||
QY 386 CAAACGGTGGATGCGAGGACTGTACAGGGGCTGTCTTAACAGCAGCGTGGTAGATAAA 445
Db |||||
QY 101 -----Gln 101
QY 446 TCTACCACTTTATAGGTTTCTTCGCTATCTTTTATTGCGTCTCTCGAAGTTCGATA 505
Db |||||
QY 102 AspPheArgPheLeuSerPheThrAspCysSerSerThrGluAspSerProSerIle 121
Db |||||
QY 506 ACTACCGGCAAGAGGCGCTTAGCTCTACGGGTAGCTTGGAGTTTGACAAAAAATGTC 565
Db |||||
QY 122 LeuTyrHisLys-----AsnGlyGlnLeuPheLeuArgAsnAsnGly 135
QY 566 AGTTTGCTCTTCAGCAAAACCTTTTCAACGGATATAGGGGTGCTTACCCCAAAACT 625
Db |||||
QY 136 AsnMetSerPheTyrArgAsnHisSerGluGlySerGlyAlaLeuSerThrAspAla 155
Db |||||
QY 626 CTTTCATTACAGGGACTACATGTCAGCTCTGTTTCTGAAATACCTCTCAAAGAAA 685
Db |||||
QY 156 LeuPheLeuGlnHisAsnTyrLeuPheThrAsnPheGluGluAsnSerSerAlaLysAsn 175
Db |||||
QY 686 GCGGAGCCATTACAGACTTCCGATGCCCTTACATTACTGGAACCAAGGGAGTCTCT 745
Db |||||
QY 176 GlyGlyAlaIleGln---AlaGlnThrLeuSerLeuSerArgAsnValSerSerLeuSer 194
QY 746 TTTTCGACAACTCTTCGGAATCTCGAGTTCGAGTCGCAATTTTACAGAACCTCGGTGACT 805
Db |||||
QY 195 PheSerArgAsnArgAlaAsnLeuAsnGlyGlyAlaIleCysCysGlnAsnLeuIleCys 214
QY 806 ATTCTTAATAATCTAAAGTTCTCTTTATTGACAAATAGTCAAGTCAAGGCGAGCTCCTCA 865
Db |||||
QY 215 SerGlyAsnValAsnProLeuPhePheThrAsnAsnSerAla----- 228
QY 866 ACAACGGGGATATGTACAGAGGTGCTATCTGTGCTTTATAAACTAGTACAGATCTAAG 925
Db |||||
QY 229 -----LeuAsnGlyGlyAlaIleCysCysIleAsnGlnAsnLeuSerGlu 244
QY 926 -----GTACCTCTACTGGAATTCAGATGTTTACTCTTCAGCAACATACATCGACA 976
Db |||||
QY 245 LysGlyCysLeuSerLeuAlaTyrAsnGlnGluThrLeuPheSerGlyAsnSerAlaLys 264
QY 977 ACAGCGGGAGGACTATCTATGTAAGAAAGCTCGAAGTGGCTTCCGAGGACTTACCCTA 1036
Db |||||
QY 265 GluLysGlyAlaIleTyrThrLysHisMetValLeuArgHisAsnGlyProValSer 284
QY 1037 TTCAGTAGAATATGTCTCAATGGAGGTACAGCTCTTAAGGTGGAGCCATAGTATCGAA 1096
Db |||||
QY 285 PheValAsnAsnSer-----AlaLysLeuGlyGlyAlaIleAlaIleGln 299
QY 1097 GATAGTGGGAATGAGTTTATCCGCCGATAGTGGTACATGCTCTTTTAGGGAATACA 1156
Db |||||
QY 300 SerGlyGlySerLeuSerIleAlaGlyGlySerValLeuPheGlnAsnAsnSer 319
QY 1157 GTCACCTCTACTCTCTCGGCAATAGAGTAGTATCGACTTAGGAACGAGTGCAGAAAG 1216
Db |||||
QY 320 CysHisPheSerAspGlnGlyThrValArgAsnAlaIleTyrLeuGluLysAsnAlaLeu 339
QY 1217 ATGACAGCTTTCCTGCTGCTGGTGGAGCCATCTACTTCTATGATGCCATAACTACA 1276
Db |||||
QY 340 LeuSerSerLeuGluAlaArgHisGly---AspIleLeuPhePheAspProIleValGln 358
QY 1277 GGA-----TCATCCACACAGTTACAGATGCTCTTAAAGTTAATGAG 1318
Db |||||
QY 359 GluValValSerProGluPheSerThrSerAlaLeuThrProLeuArgIle---Gln 377
QY 1319 ACTCCGCGCAGATTCTCGACTACATATACAGGGAACATCTCTCACAGGAGAAAGTTA 1378
Db |||||
QY 378 ThrAsnThrAsnArgAla-----ValIlePheSerSerGluAsnLeu 391

QY 1379 TCAGAGACAGGCGCGAGATTCTTAAATACTTACTTGAAGCTACTACAGCCTGTAACT 1438
Db |||||
QY 392 SerLysGluGluLysThrGluAla---AsnLeuIleSerLysIleGlnProIleGlu 410
QY 1439 CTTTCAGGAGTACTCTATCTTTTAAACATCGAGTGAATCTTCAGAGCTTCAGGCATTCAC 1498
Db |||||
QY 411 LeuGlnSerGlyCysLeuValLeuLysAspArgValIleLeuSerAlaProSerLeuSer 430
QY 1499 CAACAGGAGATCTCTGCTCGAAATGAGACTAGGAATCTCTAGAACCTCTGATGACT 1558
Db |||||
QY 431 GlnAlaProGlnAlaLeuValMetAspValGlyThrSerLeuThrSerSerAsp 450
QY 1559 AGCACCATAACAATTTGCTCATTAACATCAGTTCTATAGACGCTGCAAGAGGCAAAA 1618
Db |||||
QY 451 LeuLysLeuThrThrLeuSerIleProLeuHisSerIleAspThrGluAsnSerValSer 470
QY 1619 ATAGAAACCAAGCTACGTC---AAAATCTGACTTTATCTGGAACCATCACTTTATTG 1675
Db |||||
QY 471 IleGlnSerProThrLeuSerIleGlnLysIlePheLeuSerAsnSer----- 486
QY 1676 GACCCGAGCGGACGCTTTTATGAAATCATAGTTTAAAGAAATCTCTAGCTCTACGACATC 1735
Db |||||
QY 487 ---GluHisGluAsnPheTyrGluAsnValGluLeuLeuSerLysAspGlnLysAspIle 505
QY 1736 TTAGAGCTCAAGCTTCTGGAACCTGTAACAAGCCGAGTCACTCCAGAT---CCTATA 1792
Db |||||
QY 506 ProLeuLeu-----SerLeuProLysGlyLeuProHisProAspLeuProAsp 521
QY 1793 ATGGGTGAGAAATTCATTACGGCTTATCAGGAACTTGGGGCCCAATTTGTTGGGGACA 1852
Db |||||
QY 522 GlyAsnLeuSerSerHisPheGlyTyrGlnGlyAspTyrPasn---PheSerTyrGlnThr 540
QY 1853 GGG---GCTTTCAGACTGCAACCTTCAACTGAGCTAAACCTGGCTATATCTCTAATCCC 1909
Db |||||
QY 541 SerAspGlnArgGluThrLeuValAlaAsnTyrThrAlaAsnSerTyrIleProHisPro 560
QY 1910 GAGCTATCGGCTCTTATAGTCCCTATAGCTTATGGAATGCACTTATAGATATTAGTCT 1969
Db |||||
QY 561 GluArgGlnSerAlaLeuValAlaAsnThrLeuTyrPasnThrTyrSerAspMetGlnAla 580
QY 1970 CTCCTATTCTTATGGAGACTGCAAAACGAAGGTTTCAGGGAGACCGTCTCTTT----- 2023
Db |||||
QY 581 ValGlnSerMetIleAsnThrThrAlaGlnGly-----GlyAlaTyrLeuPheGlyThr 598
QY 2024 TGGTGTGCTGGAATTAATCTAATCTTCTC-----CATAGGATAGTACAAAACACAGCC 2077
Db |||||
QY 599 TrpGlySerAlaValSerAsnLeuPheTyrSerHisGlyAsnSerGlyLysSerThrAsp 618
QY 2078 GGGTTTCGCAATTCAGTCGCGGTTATGTCATAGGAGGAAACCTACATACTTGTTCAGAT 2137
Db |||||
QY 619 AsnTrpLysHisArgSerLeuGlyTyrLeuPheGlyIleSerThrHisSerLeuAspAsp 638
QY 2138 AAGATTCTTAGTCTCTTTTGTCAAGCTCTTTTGGAGAGATAGAGACTACTTTGTAGCT 2197
Db |||||
QY 639 HisSerPheCysLeuAlaAlaGlyGlnLeuPheGlyLysSerSerAspSerPheValThr 658
QY 2198 AGAATCAAGGTACGCTACGAGGAACTCTCTATTACCAGCAACAGAAACCTATATC 2257
Db |||||
QY 659 SerAlaAspThrThrSerTyr----- 665
QY 2258 TCTCTTCTTGGAAACTACGGCTTGTCTGTTCTTATGTTCTTACAGAGATTCCT--- 2314
Db |||||
QY 666 -----IleAlaIleGlnThrGlnIleAlaThr 675
QY 2315 -----GTTCTCTTTTTCAGAAACCTTAGCTACCCCATACGATACGATCTTGAACCC 2368
Db |||||
QY 676 SerLeuIleLysIleSerAlaGlnAlaCysTyrAsnGluSerIleHisGluLeuLysThr 695
QY 2369 AAGTATACAACTATCTCTGTTAA---GGAGCTGGGGGAATGATAGTTTCGCTTTA 2425
Db |||||
QY 696 LysTyrArgSerPheSerLysGluGlyPheGlyAlaTyrPheHisSerValAlaValSerGly 715
QY 2426 GAATTCCGTGGAAGAGCTCCGATTCTTATAGATGAAGTCTCTATTGAGCAGTACATG 2485

Db	716	GlulilecylAlaSerIleProIleValSerAsnGlySerGlyLeuPheSerPheSer	735
QY	2486	CCCTTCATGAAATTCAGATTGTCATGACATCAGCAAGCGTGTAAAGAACAGGGAACA	2545
Db	736	IlePheSerLysLeuGlnGlyPheSerGlyGlnAspGlyPheGlnGluSerArgGly	755
QY	2546	GAAGCTCGTGAATTTGGAAGTAGCCGCTTGTGAATCTTCCTACCTCGGGATCCGA	2605
Db	756	GluAlaArgAlaPheAlaAspSerSerPheThrAsnIleSerLeuProValGlyIleAla	775
QY	2606	TTTGATGAAGAAATCAGACTGCCAAGATGCACGTAACAATCTCTGCTTATCTGTG	2665
Db	776	PheGluLysLysSerGlnLysThrArgAsnTyrTyrHisPheLeuGlyAlaTyrIleGln	795
QY	2666	GATCTTCTGCTAGTAACCCCGACTGACGACAACTGCGAATTAACGCGTGATCTTGG	2725
Db	796	AspLeuLysArgCysValGluSerGlyProValThrLeuLeuLysAsnSerValThrTrp	815
QY	2726	AAACCTTCCTCGTAGCAATTTGGCAAGACAAGCTTTAGTCCTCTCGTCAGGAACCAATTTT	2785
Db	816	AspAlaProMetAlaAlaAsnLeuAspSerArgAlaTrpMetPheArgLeuThrAsnGlnArg	835
QY	2786	TGCTTTAACTCAAAATTTGAAGCCTTTAGCCAAATTTCTTTGAAATTCGCTGGGTCACT	2845
Db	836	AlaLeuHis---ArgPheGlnThrLeuValAsnMetSerTyrMetLeuArgGlyGlnSer	854
QY	2846	CGCAATTTACAATGTAGACTTAGGAGCAAAATACCAATTC	2884
Db	855	TyrSerTyrSerLeuAspLeuGlyThrThrTyrArgPhe	867
RESULT 13			
PMP1 CHLTR			
ID	PMP1	CHLTR	STANDARD; PRT; 878 AA.
AC	O84882;		
DT	16-OCT-2001	(Rel. 40, Created)	
DT	16-OCT-2001	(Rel. 40, last sequence update)	
DT	16-OCT-2001	(Rel. 40, last annotation update)	
DE	Probable outer membrane protein pmp1 precursor (Polymorphic membrane protein 1).		
DE	PMP1 OR CT874.		
OS	Chlamydia trachomatis.		
OC	Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.		
OC	NCBI_TaxID=813;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=D/UK-3/Cx.		
RX	MEDLINE=9900809; PubMed=9784136;		
RA	Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,		
RA	Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,		
RA	Davis R.W.;		
RT	"Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis";		
RL	Science 282:754-759(1998).		
CC	-!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES) (POTENTIAL).		
CC	-!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
DR	EMBL;	AE001361;	AAC68472.1; -.
DR	PIR;	B71460;	B71460.
DR	InterPro;	IPR003368;	Chlamydia_PMP.
DR	Pfam;	PF02415;	DUF145; 1
DR	TIGRFAMs;	TIGR01376;	POMP repeat; 6.
DR	outer membrane; Signal; Multigene family; Complete proteome.		
FT	SIGNAL		
FW	1 24		

```

QY 1274 ACAGGATCATCCACAACAGTTACAGATGTCTTAAAGTAAATGAGACTCCGGCAGATTCT 1333
Db      |||||:::
QY 357 GlnGluSerSerSer-----LysGluSerProLeuProSer 368
QY 1334 GCACCTACATATACAC----- 1348
Db      |||||:::
QY 369 SerLeuGlnAlaSerValThrSerProThrProAlaThrAlaSerProLeuValIleGln 388
QY 1349 -----GGGAACATCATCTTCAGAGAGAAAGTATATACAGACAGAGCCGCA 1396
Db      |||||:::
QY 389 ThrSerAlaAsnArgSerValIlePheSerSerGluArgLeuSerGluGluGluLysThr 408
QY 1397 GATTCTAAAATCTTACTCGAGCTACTACAGCTGTAACTCTTTCAGGAGGTACTCTA 1456
Db      |||||:::
QY 409 ProAsp---AsnLeuThrSerGlnLeuGlnGlnProIleGluLeuLysSerGlyArgLeu 427
QY 1457 TCTTTAAACATCGGAGTCTCGCAGACTCGAGCATTCACCTCAACAGCAGCATCTCGT 1516
Db      |||||:::
QY 428 ValLeuLysAspArgAlaValLeuSerAlaProSerLeuSerGlnAspProGlnAlaLeu 447
QY 1517 CTCGAAATGAGCGTAGCACTACTCTAGAACCTGTGATCTAGACACCAATAACAATTG 1576
Db      |||||:::
QY 448 LeuIleMetGluAlaGlyThrSerLeuLysThrSerSerAspLeuLysLeuAlaThrLeu 467
QY 1577 GTCATTAAACATCAGTCTATAGACGCTGCAAGAGCGCAAAATAAGAACCAAGCTACG 1636
Db      |||||:::
QY 468 SerIleProLeuHisSerLeuAspThrGluLysSerValThrIleHis----- 483
QY 1637 TCAAAAAATCTGACTTTATCTGGAACCATCATTATTGGACCCGCGGCGCTTTAT 1696
Db      |||||:::
QY 484 AlaProAsnLeuSerIleGlnLysIlePheLeuSerAsnSerGlyAspGluAsnPheTyr 503
QY 1697 GAAATCATAGTTTAAAGAAATCTCAGTCTAGCACTCTAGACATCTTAGAGCTCAAGCTTCTGA 1756
Db      |||||:::
QY 504 GluAsnValGluLeuSerLysGluGlnAsnAsnIleProLeuLeuThrLeuSer--- 522
QY 1757 ACTGTAAACAGCACCGCAGTACTCCAGATCTCTATATGGTGAGAAATTCATTACGGC 1816
Db      |||||:::
QY 523 LysGluGlnSerHisLeuHisLeuProAspGlyAsnLeuSerSer-----HisPheGly 540
QY 1817 TATCAGGGAACCTTGGGCGCCCAATTGTTTG-----GGGACAGGGGCTTACGACT 1867
Db      |||||:::
QY 541 TyrGlnGlyAspTrp---ThrPheSerTrpLysAspSerAspGluGlyHisSerLeuIle 559
QY 1868 GCAACCTTCACTGAGCTAAACATCGGTATATCTTAATCCAGAGCGGTATCGCTCTTA 1927
Db      |||||:::
QY 560 Ala-----AsnTrpThrProLysAsnTyrValProHisProGluArgGlnSerThrLeu 577
QY 1928 GTCCTTAATAGCTTATGGAATGCATTTATAGATATTAGCTCTCTCCATTATCTATGAG 1987
Db      |||||:::
QY 578 ValAlaAsnThrLeuTrpAsnThrTyrSerAspMetGlnAlaValGlnSerMetIleAsn 597
QY 1988 ACTGCAAAAGAGGTTGACGGGAGACCGTGCTTTT-----TGGTGCTGCTGATTATCT 2041
Db      |||||:::
QY 598 ThrIleAlaHisGly-----GlyAlaTyrLeuPheGlyThrTrpGlySerAlaValSer 615
QY 2042 AACTCTTCT-----CATAGGATAGTACAAAACAGCAGCGCGGTTCGCCATTGAGT 2095
Db      |||||:::
QY 616 AsnLeuPheTyrAlaHisAspSerSerGlyLysProIleAspAsnTrpHisHisArgSer 635
QY 2096 GCGGTTTATGTCATAGGAGGAACCTACATCTTGTTCAGATAAGATTCTTAGTGCTGCA 2155
Db      |||||:::
QY 636 LeuGlyTyrLeuPheGlyIleSerThrHisSerLeuAspHisSerPheCysLeuAla 655
QY 2156 TTTTGTGAGCTCTTTGGAAGAGATAGAGACTTCTTTGTAGCTAAGATCAAGGTACAGTC 2215
Db      |||||:::
QY 656 AlaGlyGlnLeuLeuGlyLysSerSerAspSerPheIleThrSerThrGluThrThrSer 675
QY 2216 TAGGAGGAACTCTCTATATACAGCAACAGAACTATATCTCTCTCTCTGCAACTA 2275
Db      |||||:::
QY 676 TyrIleAlaThrValGlnAlaGln-----LeuAlaThrProLeu----- 688
QY 2276 CGGCCTTGTCTGTCTTATGTTCTCTACAGAGATTCTGTTCTCTTTTCAGGAACCTT 2335

```

```

Db      |||||:::
QY 689 -----MetLysIleSerAlaGlnAla 695
QY 2336 AGCTACACCCATACGGATACGATCTGAAACCAAGTATATACACATATCTTACTGTTAA 2395
Db      |||||:::
QY 696 CysTyrAsnGluSerIleHisGluLeuLysThrLysTyrArgSerPheSerLysGluGly 715
QY 2396 ---GGAAGCTCGGGAATGATAGTTTGCCTTTAGAAATTCGGTGAAGAGCTCCGATTGC 2452
Db      |||||:::
QY 716 PheGlySerTrpHisSerValAlaValSerGlyGluValCysAlaSerIleProIleVal 735
QY 2453 TTAGATGAAGTCTCTATTTCAGCAGTACATCCCTTCATGAAATTCGATTTGTCTAT 2512
Db      |||||:::
QY 736 SerAsnGlySerGlyLeuPheSerSerPheSerIlePheSerLysLeuGlnGlyPheSer 755
QY 2513 GCACATCAGGAAGGTTTAAAGAACAGGACAGAGCTCGTGAATTTGGAGTAGCCGT 2572
Db      |||||:::
QY 756 GlyThrGlnAspGlyPheGluLysSerGlyGluIleArgSerPheSerAlaSerSer 775
QY 2573 CTTGTGAATCTTGCCTTACCTATCGGATCCGATTTTGATAAGGAATCAGACTGCCAAGAT 2632
Db      |||||:::
QY 776 PheArgAsnIleSerLeuProMetGlyIleThrPheGluLysLysSerGlnLysThrArg 795
QY 2633 GCAACGTACAACTTAACTCTTGGTATATCTGTGGATCTTGTTCGAGTAACCCCGACTGT 2692
Db      |||||:::
QY 796 AsnTyrTyrThrPheLeuGlyAlaTyrIleGlnAspLeuLysArgAspValIleSerGly 815
QY 2693 ACGCAACACACTCGCAATTAGCGGTGATTTCTTGGAAACCTTCGGTACCAATTTGCAAGA 2752
Db      |||||:::
QY 816 ProValValLeuLeuLysAsnAlaValSerTrpAspAlaProMetAlaAsnLeuAspSer 835
QY 2753 CAAGCTTTAGTCTCTTGGTGCAGGGAACCATTTTCTTTAACTCAATTTTGAAGCCTTT 2812
Db      |||||:::
QY 836 ArgAlaTyrMetPheArgLeuThrAsnGlnArgAlaLeuHis---ArgLeuGlnThrLeu 854
QY 2813 ACCCAATTTCTTTGAAATTCGTCGTCATCTCCCAATTACCAATTAGACTTAGAGCA 2872
Db      |||||:::
QY 855 LeuAsnValSerTyrValLeuArgGlyGlnSerHisSerTyrSerLeuAspLeuGlyThr 874
QY 2873 AAATACCAATTC 2884
Db      |||||:::
QY 875 ThrTyrArgPhe 878

RESULT 14
PM21 CHLPN
ID PM21_CHLPN STANDARD; PRT; 1609 AA.
AC Q9Z6U5; Q9RB58;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable outer membrane protein pmp21 precursor (Polymorphic membrane
DE protein 21).
OS PMP21 OR CPN0963 OR CP0897.
GN Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Praser C.M.;

```

RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
 RL pneumoniae AR39.";
 RN Nucleic Acids Res. 28:1397-1406(2000).
 RP [3]
 RC SEQUENCE FROM N.A.
 RX STRAIN-J138.
 RC MEDLINE=20330349; PubMed=10871362;
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 from Japan and CWL029 from USA.";
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -1- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
 CC (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AE001676; AAD19099.1; -;
 DR EMBL; AE002248; AAF38684.1; -;
 DR EMBL; AP002548; BAA99171.1; -;
 DR PIR; A86611; A86611.
 DR PIR; H72013; H72013.
 DR PHCI-2DPAGE; Q9RB58; -;
 DR TIGR; CP0897; -;
 DR InterPro; IPR006315; Autotransport.
 DR InterPro; IPR005546; Autotransporter.
 DR InterPro; IPR003368; Chlamydia PMP.
 DR Pfam; PF03797; Autotransporter; 1.
 DR Pfam; PF02415; DUF145; 1.
 DR TIGRFAMs; TIGR01414; autotrans_bar1; 1.
 DR TIGRFAMs; TIGR01376; POMP_repeat; 13.
 KW Outer membrane; Signal; Multigene family; Complete proteome.
 FT SIGNAL 1 30
 FT CHAIN 31 1609 PROBABLE OUTER MEMBRANE PROTEIN PMP21.
 FT CONFLICT 420 420 I -> M (IN REF. 3).
 SQ SEQUENCE 1609 AA; 170865 MW; 2604C3E9FC4024CB CRC64;

 Alignment Scores:
 Pred. No.: 7,05e-34 Length: 1609
 Score: 685.50 Matches: 254
 Percent Similarity: 38.91% Conservative: 144
 Best Local Similarity: 24.81% Mismatches: 372
 Query Match: 12.67% Indels: 253
 DB: 1 Gaps: 36

 US-09-428-122-1 (1-3000) x PM21_CHLPN (1-1609)
 QY 278 GTCACCTAGAAAATATTCCTGGACAGGACGCAATACACAAAAGCTGTTT----- 331
 Db ValSerSerThrAspIleArgGlyGlyAlaIleLeuAlaGlnHisIlePheThr 705
 QY 332 AACACACTAAGGCGATTGACTTTCACAGGTACGGGAATCTCTATTGTTCCAAACG 391
 Db AspAsnThr---GlyAsnLeuArgPheSerGlyAsn----- 716
 QY 392 GTGGATGACGAGGACTGTAGCGGGCTGCTGTAAACAGACGCGTGGTAAATCTACC 451
 Db -----LeuGlyGlyGlyGluGluSerThrVal----- 726
 QY 452 ACCTTTATAGGTTTCTTCGCTATCTTTTATTGCTCTCTGGAGTTCGTAATACTACC 511
 Db -----GlyAspLeuAlaIleVal 732
 QY 512 GGCAAGGACCGGTAGCTGCTACGGGTAGCTTGGTTTGACAAAATGTCAGTTTG 571
 Db -----LeuSerThrAsnGluValAsnValCysSerAsnGlnAsnVal 751

QY 572 CTCTTCAGCAAAACATTTTCAACG-----GATAATGGCGTCTATCACCAGAAA 622
 Db ValPheSerAspAsnValThrSerAsnGlyCysAspSerGlyAlaIleLeuAlaLys 771
 QY 623 ACTCTTTTATTAACAGGACTACAAATGTCTGCTCTCTGTTTCTGAAATATCTCTCAAAG 682
 Db LysValAspIleSer---AlaAsnHisSerValGluPheValSerAsnGlySerGlyLys 790
 QY 683 AAGGCGGAGCCATT---CAGACTTCGATGCCCTTACCATTTACTGGAACACAGGGGAA 739
 Db PheGlyGlyAlaValCysAlaLeuAsnGluSerValAsnIleThrAspAsnGlySerAla 810
 QY 740 GTCTCTTTTCTGACATATCTCTCGGATTCGGAGCTGCAATTTTACA---GAAGCC 796
 Db ValSerPheSerLysAsnArgThrArgLeuGlyGlyAlaGlyValAlaAlaProGlnGly 830
 QY 797 TCGGTGACTATT---TCTAATAATGCTAAAGTTTCTCTTTTATTGACATTAAGTGCACAGGA 853
 Db SerValThrIleCysGlyAsnGlnGlyAsnIleAlaPheLysGluAsnPheValPheGly 850
 QY 854 GCGAGC----- 859
 Db SerGluAsnGlnArgSerGlyGlyAlaIleAlaAsnSerSerValAsnIleGln 870
 QY 860 -----TCCTCAACACGGGGATATGTACAGG----- 886
 Db AspAsnAlaGlyAspIleLeuPheValSerAsnSerThrGlySerThrGlyGlyAlaIle 890
 QY 887 -----GGTCTATCTGTCTATATAAAGTACATACATAGTAAAGTCCACCTCAGTGA 940
 Db PheValGlySerLeuValAlaSerGluGlySerAsnProArgThrLeuThrIleThrGly 910
 QY 941 AATCAG---ANGTTACTCTTCAGCAACAATACATACGACACACAGCG----- 982
 Db AsnSerGlyAspIleLeuPheAlaLysAsnSerThrGlnThrAlaAlaSerLeuSerGlu 930
 QY 983 -----GGAGGAGCTATCTATGTGAAAAGCTCGAACTGGCT-----TCC 1021
 Db LysAspSerPheGlyGlyAlaIleThrGlnAsnLeuLysIleValLysAsnAla 950
 QY 1022 GGAGGACTTACCTTATTCAGTAGAATAGTGTCAATGAGGATGACAGTCTCTTAAAGTGA 1081
 Db GlyAsnValSerPheTyr-----GlyAsnArgAlaProSerGlyAla 964
 QY 1082 GCCATAGCTATCGAAGATAGTGGGAATGAGTTTATCCGCCGATAGTGTGACATTGTC 1141
 Db GlyValGlnIleAlaAspGlyGlyThrValCysLeuGluAlaPheGlyGlyAspIleLeu 984
 QY 1142 TTTTATAGGATACAGTCACTTCTACTCTCTGGAGCAATAGAGTAGTAGTATCGACTTA 1201
 Db PheGluGlyAsnIle-----AsnPheAspGlySerPheAsnAlaIleHisLeuCys 1001
 QY 1202 GGAACGAGTGCAAAGATGACAGCTTTGCGTTTCTGCTGCTGATAGAGCATCTACTCTAT 1261
 Db GlyAsnAspSerLysIleValGluLeuSerAlaValGlnAspLysAsnIleIlePheGln 1021
 QY 1262 GATCCCATTAATCAGGATCATCCACA-----ACAGTTACAGTGC----- 1303
 Db AspAlaIleThrThrGluAsnThrIleArgGlyLeuProAspLysAspValSerPro 1041
 QY 1304 -----TTAAAAGTTAATGACACTCCGGCA---GATTCTGCCTACATAAT 1345
 Db LeuSerAlaProSerLeuIlePheAsnSerLysProGlnAspAspSerAlaGlnHis 1061
 QY 1346 ACAGGCAACATCATCTTCACAGGAGAAAAGTTATCAGACAGAGCGCGCAGATTCTAAA 1405
 Db GluGlyThrIleArgPheSer----- 1068
 QY 1406 AATCTTACTCGAAGCTACTACGCTGTAACTCTTTTCAGGAGGTACTCTATCTTTTAAA 1465
 Db ArgGlyValSerLysIleProGlnIleAlaAlaIleGlnGluGlyThrLeuAlaLeuSer 1089

```
QY 1466 CATGGAGTCTTCGAGACTCAGGCATTCACTCAACAGGAGATTCCTGCTCGAAATG 1525
Db 1089 GlnAsnAlaGluLeuTTPLeuAlaGlyLeuLysGlnGluThrGlySerSerIleValLeu 1108
QY 1526 GACGTAGAACTACTCTAGAACCTGCTGATACCTAGCACCATAACAAT----- 1573
Db 1109 SerAlaGlySerIleLeuArgIlePheAspSerGlnValAspSerSerAlaProLeuPro 1128
QY 1574 -----TTGGTCAATAACATCAGTTCTATA 1597
Db 1129 ThrGluAsnLysGluGluThrLeuValSerAlaGlyValGlnIleAsnMetSerSerPro 1148
QY 1598 GACGGTCAAGAGAGGCAAAATAGAACCAAGCTAGCTCAAAATCTGACATTATCT 1657
Db 1149 ThrProAsnLysAspLysAlaValAspThrProValLeuAlaAspIleIleSerIleThr 1168
QY 1658 -----GSAACCACTCACTTTA----- 1672
Db 1169 ValAspLeuSerSerPheValProGluGlnAspGlyThrLeuProLeuProGluIle 1188
QY 1673 -----TTGGACCCG 1681
Db 1189 IleIleProLysGlyThrLysLeuHisSerAsnAlaIleAspLeuLysIleIleAspPro 1208
QY 1682 ACGGCACGTTTTATGAAATCATAGT---TTAAGAAATCTCTCAGTCTCAGACATCTTA 1738
Db 1209 ThrAsnValGlyTyrGluAsnHisAlaLeuLeuSerSerHisLysAspIleProLeuIle 1228
QY 1739 GAGCTCAAGCTTCGGAAGCTGTAACAAGCACCGCAGTGACTCCAGATCTTATATGGGT 1798
Db 1229 SerLeuLysThrAlaGluGlyMetThrGlyThrProThrAlaAspAlaSerLeuSerAsn 1248
QY 1799 GAGAAATCCAT-----TACGGTATCAGGGAAT 1828
Db 1249 IleLysIleAspValSerLeuProSerIleThrProAlaThrTyrGlyHisThrGlyVal 1268
QY 1829 TGG-----GGCCAAATGTTGGGGACAGGGCTCTCAGACT 1867
Db 1269 TrpSerGluSerLysMetGluAspGlyArgLeuValValGly----- 1282
QY 1868 GCAACCTTCACTGACTAAATCGCTATATTCCTAATCCGAGCGTATCGGCTCTTTA 1927
Db 1283 -----TrpGlnProThrGlyTyrLysLeuAsnProGluLysGlnGlyAlaLeu 1298
QY 1928 GTCCCTAATPAGCTTATGGAATGCAATTATAGATATTAGCTCTCTC----- 1972
Db 1299 ValLeuAsnAsnLeuTrpSerHisTyrThrAspLeuArgAlaLeuLysGlnGluIlePhe 1318
QY 1973 ----CATTATCTTATGAGACTGCAACGAGGGTTCACGGAGACCGTCTTTTGGTGT 2029
Db 1319 AlaHisHisThrIleAlaGlnArgMetGlu---LeuAspPheSerThrAsnValTrpGly 1337
QY 2030 GCTGGATTATCTAACTTCTTCCATAAGGATAGTACAAAAACACGACGCGGTTTCGCCAT 2089
Db 1338 SerGlyLeuGlyValValGluAspCysGlnAsnIleGlyLeuPheAspGlyPheLysHis 1357
QY 2090 TTGAGTGGCGGTATGTCATAGAGGAACCTACATACCTTGTTCAGATAGATCTTAGT 2149
Db 1358 HisLeuThrGlyTyrAlaLeuGlyLeuAspThrGlnLeuValGluAspPheLeuIleGly 1377
QY 2150 GCTGCATTTTGTGAGCTCTTTTGAAGA---GATAGAGACTACTTTCTAGCTAAGAT--- 2203
Db 1378 GlyCysPheSerGlnPheGlyLysThrGluSerGlnSerTyrLysAlaLysAsnAsp 1397
QY 2204 -----CAAGGTACAGTCTACGGAGGAGTCTCTATTACCAGCACACGAAC 2251
Db 1398 ValLysSerTyrMetGlyAlaIleTyrAlaGlyIleLeu----- 1410
QY 2252 TATATCTCTCTCTCTGCAAACTACGGCCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2311
Db 1411 -----AlaGly 1412
QY 2312 CCTGTTCTCTTTTCAGGAAACCTTAGCTACACCCATACCGATAACGATCTGAAACCAAG 2371
```

```
Db 1413 ProTrpLeuIleLysGlyAlaPheValTyrGlyAsnIleAsnAspLeuThrThrAsp 1432
QY 2372 TATACACATPATCTACTGTATAAGAGAGCTGGGGAATGATAGTTTCGCTTTAGAAATC 2431
Db 1433 TyrGlyThrLeuGlyIleSerThrGlySerTrpIleGlyLysGlyPheIle----- 1449
QY 2432 GGTGGAAGAGCTCGATTTCCTTAGATGA----- 2461
Db 1450 -----AlaGlyThrSerIleAspTyrArgTyrIleValAsnProArgArgPheIle 1466
QY 2462 AGTGCTCTATTGAGCAGTACATCCCTTCATGAATTCGAGTTTGTCTATGCACATCAG 2521
Db 1467 SerAlaIleValSerThrValValProPheValGluAlaGluTyrValAlaGlyAspLeu 1486
QY 2522 GAAGTTTAAAGAACAGGAGACAGAGCTCGTGAATTTGGAAGTACCGCTCTGTGAAT 2581
Db 1487 ProGluIleSerGluGlnGlyLysGluValArgThrPheGlnLysThrArgPheGluAsn 1506
QY 2582 CTTGCCCTTACTATCGGGATCCGATTGATAAG-----GAATCAGACTGCCAAGATGCA 2635
Db 1507 ValAlaIleProPheGlyPheAlaLeuGluHisAlaTyrSerArgGlySerArgAlaGlu 1526
QY 2636 ACGTACATTAATCTTGTGTTTACTGTGATCTTGTTCGTAGTAAACCCGACTGTACG 2695
Db 1527 ValAsnSerValGlnLeuAlaTyrValPheAspValTyrArgLysGlyProValSerLeu 1546
QY 2696 ACAACACTGCAATAGCGGTGATCTTGGAAACCTTCGTAGCAATTTGGCAAGACAA 2755
Db 1547 IleThrLeuLysAspAlaAlaTyrSerTrpLysSerTyrGlyValAspIleProCysLys 1566
QY 2756 GCTTTAGTCTCTCGTCAGGAGAACATTTTGTCTTAACTCAAAATTTTGAAGCCTTTAGC 2815
Db 1567 AlaTrpLysAlaArgLeuSerAsnAsnThrGluTrpAsnSerTyrLeuSerThrTyrLeu 1586
QY 2816 CAATTTCTTTGAATTCGTCGTCGTCATCTCGCAATTAACAATGATAGACTTAGGAGCAAA 2875
Db 1587 AlaPheAsnTyrGluTrpArgGluAspLeuIleAlaTyrAspPheAsnGlyGlyLearg 1606
QY 2876 TACCNAATC 2884
Db 1607 IleIlePhe 1609

RESULT 15
PM14 CHLPN CHLPN STANDARD; PRT; 978 AA.
ID PM14 CHLPN STANDARD; PRT; 978 AA.
AC Q92895; Q9RB63;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable outer membrane protein pmp14 precursor (Polymorphic membrane
DE protein 14).
GN PMP14 OR CPN0454 OR CP0298.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI TaxID=83558;
[1] _SEQUENCE FROM N.A.
RN _SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Bisen J., Fraser C.M.;
RN [2]
RN _SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Bisen J., Fraser C.M.;
```

"Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39." ; 28:1397-1406 (2000).

RT Nucleic Acids Res. 28:1397-1406 (2000).

RL [3]

RN SEQUENCE FROM N.A.

RP STRAIN=J138;

RC MEDLINE=20330349; PubMed=10871362;

RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K., Shiba T., Ishii K., Hattori M., Kunara S., Nakazawa T.;

RA "Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CW1029 from USA." ;

RT Nucleic Acids Res. 28:2311-2314 (2000).

RL CC -1- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)

CC -1- (POTENTIAL). BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

CC EMBL; AE001629; AAD18596.1; -

DR EMBL; AE002191; AAF38155.1; -

DR EMBL; AP002546; BAA98661.1; -

DR PIR; B81593; B81593.

DR PIR; C86547; C86547.

DR PIR; G72076; G72076.

DR PHCI-2DPAGE; Q92895; -

DR TIGR; CP0298; -

DR InterPro; IPR001368; Chlamydia_pmp.

DR Pfam; PF02415; DUF145.1

DR TIGRFAMS; TIGR01376; POMP repeat; 6.

KW Outer membrane; Signal; Multigene family; Complete proteome.

FT SIGNAL 1 24 POTENTIAL.

FT CHAIN 25 978 PROBABLE OUTER MEMBRANE PROTEIN PMP14.

FT CONFLICT 379 379 N -> D (IN REF. 1).

FT SEQUENCE 978 AA; 103654 MW; 6EEE1429999D3019 CRC64;

SQ

Alignment Scores:

Pred. No.: 7.8e-33 Length: 978

Score: 667.50 Matches: 272

Percent Similarity: 39.11% Conservative: 134

Best Local Similarity: 26.20% Mismatches: 445

Query Match: 12.34% Indels: 187

DB: 1 Gaps: 37

US-09-428-122-1 (1-3000) x PM14_CHLPN (1-978)

Qy 149 CTTTGTCTATGTTGTCACGAGACAGTTTTC-----GATTCAAGTGGC 193

Db 2 ProLeuSerPheLysSerSerPheCysLeuLeuAlaCysLeuSerAlaSerCys 21

Qy 194 AGTTTC---GATGGGAATAAAATGTTATTTTCAGTTCGTGAGTACGGAAGTGT 250

Db 22 AlaPheAlaGluThrArgLeuGlyGlyAsnPheValProPheIleThrAsnGln----- 39

Qy 251 GGAACCTACCTATTATTAAAGGAATGTCACTCTAGAAAAATTCCTCGAACAGGCACA 310

Db 40 GlyGluLeuLeuLeuThrSerAspPheValCysSerAsnPheLeuGly----- 56

Qy 311 GCAATCAAAAAGCTGTTTAAACACATAAGGGGATTTGACTTTCACAGTAACGGG 370

Db 57 AlaSerPheSerPheLeuAsnSerSerSerAsnLeuSerLeuLeuGlyLysGly 76

Qy 371 AACTCTCTATTGTTCCAAACGGTGGATGAGGAGCTGTAGCAGGGCTGCTGTTAACAGC 430

Db 77 LeuSerLeuThrPheThrSerCysGlnAlaProThrAsnSerAsnTyrAlaLeuLeuSer 96

Qy 431 AGCGTGGTAGATAAATCTACACCGCTTTATAGGGTTTTTCTTCGCTATCTTTATTGCGTCT 490

Qy


```
QY 1397 GATTCTAAAATCTTACTTCGAAGCTACTACAGCCTGTAACTCTTTTCAGGAGGTACTCTA 1456
Db 461 LeuProAlaAsnThrThrIleLeuLeuGlyThrValLysIleAlaSerGlyGluLeu 480
QY 1457 TCTTTAAACATGAGAGTACTCTCGAGACTCAGGCATTCCTCAACAGCGAGATTCCTCGT 1516
Db 481 LysIleThrAspAsnAlaValValAsnValLeuGlyPheAlaThrGlnGlySerGlyGln 500
QY 1517 CTCGAAATGGACGTAGGACTACTCTA-----GAACCTCGCTGAT 1555
Db 501 LeuThrLeuGlySerGlyGlyThrLeuGlyLeuAlaThrProThrGlyAlaProAlaAla 520
QY 1556 ACTAGC---ACCATAAACAAATTGGTCATTAAACATCAGTTCTATAGACGGTGCAGAAAG 1612
Db 521 ValAspPheThrIleGlyLysLeuAlaPheAspProPheSerPheLeuLysArgAspPhe 540
QY 1613 GCAAAATAGAAACCAAGCTACGTCAAAAATCTGACTTTATCTGGAACCATCACTTTA 1672
Db 541 ValSerAlaSerValAsnAlaGlyThrLysAsnValThrLeuThrGlyAlaLeuValLeu 560
QY 1673 TTGGACCCGACGGCAGC---TTTTATGAAATCATAGTTTAAAGAAATCCTCAGTCTCTAC 1729
Db 561 AspGluHisAspValThrAspLeuTyrAspMetValSerLeuGlnSerProValAlaIle 580
QY 1730 GACATCTTAGAGCTCAAGCTTCTGGAACGTGTAACAGCACCGCAGTGACTCCAGATCCT 1789
Db 581 ProfileAlaValPheLysGlyAla---ThrValThrLysThrGlyPhe---ProAspGly 598
QY 1790 ATAATGGGTGAGAAATTCATTACGGCTATCAGGNACTGG-----1831
Db 599 GluIleAlaThrProSerHisTyrGlyTyrGlnGlyLysTyrSerTyrThrTrpSerArg 618
QY 1832 -----GGCCCA-----1837
Db 619 ProLeuLeuIleProAlaProAspGlyGlyPheProGlyGlyProSerProSerAlaAsn 638
QY 1838 -----ATTGTTTGGGGACAGGGCTTACGACTGCAACCTTCACTGGACTAAA 1888
Db 639 ThrLeuTyrAlaValTrpAsnSerAspThrLeuValArgSerThr-----653
QY 1889 ACTGGCTATATTCCTAATCCGAGCGTATCGGCTCTTTAGTCCCTAATAGCTATGCG---1945
Db 654 -----TyrIleLeuAspProGluArgTyrGlyGluIleValSerAsnSerLeuTrpIle 671
QY 1946 -----AATGCATTTATAGATATAGCTCTCCATATTCTTATGAGAGACT 1990
Db 672 SerPheLeuGlyAsnGlnAlaPheSerAspIle-----LeuGlnAspValLeuLeuIle 689
QY 1991 GCAACGAAAGGGTTGAGGGAGACCGGTCTTTTGTGTGCTGCTGATTAICTAACTCTTC 2050
Db 690 AspHisProGlyLeu-----SerIleThrAlaLysAlaLeuGlyAlaTyrVal 705
QY 2051 CATAAGGATAGTACAAAACACGACGCGGTTTCGCCATTGTAGTGGCGGTATGTCATA 2110
Db 706 GluHisThrProArgGlnGlyHisGluGlyPheSerGlyArgTyrGlyGlyTyrGlnAla 725
QY 2111 GGAGGAAACCTACATCTTCTCAGATAAGATTTCTTAGTGTGCTGCTATTTCTCAGCTCTTT 2170
Db 726 AlaLeuSerMetAsnTyrThrAspHisThrLeuGlyLeuSerPheGlyGlnLeuTyr 745
QY 2171 GGAAGAGATAGAGACTACTTTGTAGCTAAGAAATCAAGGTACAGTCTACGGAGAACTCTC 2230
Db 746 GlyLysThr-----748
QY 2231 TATTACGACGACACAGCAACCTATATCTCTCTTCTGCAAACTACGCGCCTTGTTCGTG 2290
Db 749 -----AsnAlaAsnProTyrAspSerArgCysSerGluGlnMetTyrLeuLeu 764
QY 2291 TCTTAT-----GTTCCCTACAGAGATTCCTGTTCTCTTTTCAGGAAAC 2332
Db 765 SerPhePheGlyGlnPheProIleValThrGlnLysSerGluAlaLeuSerTrpLys 784
```

```
QY 2333 CTTAGCTACACCCATACGATAACGATCTGAAAACCAAGTAT-----ACAACATAT 2383
Db 785 AlaAlaTyrGlyTyrSerLysAsnHisLeuAsnThrThrTyrLeuArgProAspLysAla 804
QY 2384 CCTACTGTTTAAAGAAAGCTGGGGAATGATAGTTTCGCTTTTGAATTCGTGGGAGAGCT 2443
Db 805 ProLysSerGlnGlyGlnTrpHisAsnAsnSerTyrTyrValLeuIleSerAlaGluHis 824
QY 2444 CCGATT-----TCCTTAGATGAAAGTGCTCTATTTTCAGCAGTAC-----ATGCCC 2488
Db 825 ProPheLeuAsnTrpCysLeuLeuThrArgProLeuAlaGlnAlaTrpAspLeuSerGly 844
QY 2489 TTCATGAAATTCGAGTTTGTCTATGCACATCAGGAAGGTTTTTAAAGAACAGGGACAGAA 2548
Db 845 PheIleSerAlaGluPheLeuGlyGlyTrpGlnSerLysPheThrGluThrGlyAspLeu 864
QY 2549 GCTCGTGAATTTGGAAGTAGCCGCTCTGTGTAATCTTCGCTTACCTATCGGATCCGA---2605
Db 865 GlnArgSerPheSerArgGlyLysGlyTyrAsnValSerLeuProIleGlyCysSerSer 884
QY 2606 -----TTTGATAAGGAATCAGACTGCCAAGATGCAACGTACAATCTAACT 2650
Db 885 GlnTrpPheThrProPheLysLys-----AlaProSerThrLeuThrIleLys 900
QY 2651 CTTGTTTACTACTGTGGATCTTGTTCGTAGTAAACCCCGACTGTACGACAACTCGCAATT 2710
Db 901 LeuAlaTyrLysProAspIleTyrArgValAsnProHisAsnIleValThrValValSer 920
QY 2711 AGCGGTGATCTTGGAAAACCTTCGTTACGAAATTTGGCAAGACAGCAAGCTTTAGTCTTCGT 2770
Db 921 AsnGlnGluSerThrSerIleSerGlyAlaAsnLeuArgArgHisGlyLeuPheValGln 940
QY 2771 GCAGGGAACCATTTTGTCTTTTAACTCAATTTTGAAGCCTTTAGCCAATTTCTTTTCAA 2830
Db 941 IleHisAspValValAspLeuThrGluAspThrGlnAlaPheLeuAsnTyrThrPheAsp 960
QY 2831 TTGCGTGGGTCACTCTCGCAATTTACAATGTAGACTTTAGGAGCAAAATCAATTC 2884
Db 961 GlyLysAsnGlyPheThrAsnHisArgValSerThrGlyLeuLysSerThrPhe 978
```

Search completed: December 16, 2003, 10:24:11

Job time : 113.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 16, 2003, 10:21:34 ; Search time 120.5 Seconds

(without alignments)
12849.092 Million cell updates/sec

Title: US-09-428-122-1

Perfect score: 5409

Sequence: 1 cgtcttacctagtagaggt.....tggttgctaaacactttc 3000

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 1661050

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2 1/USPTO spoel p/US09428122/runat 16122003 102128 10777/app_query.fasta_1.3143
-DB=SPTREMBL 23 -OPMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09428122 @CGN 1 1 161 @runat 16122003 102128 10777 -NCPU=6 -ICPU=3
-NO MWAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTREMBL_23:*
1: sp archaea:*
2: sp bacteria:*
3: sp fungi:*
4: sp human:*
5: sp invertebrate:*
6: sp mammal:*
7: sp mhc:*
8: sp organelle:*
9: sp phage:*
10: sp plant:*
11: sp rodent:*
12: sp virus:*
13: sp vertebrate:*
14: sp unclassified:*
15: sp rvirus:*
16: sp bacteriap:*
17: sp archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3991	73.8	772	16 Q9RB71	Q9RB71 chlamydia p

ALIGNMENTS

RESULT 1

ID	Q9RB71	PRELIMINARY;	PRT;	772 AA.
AC	Q9RB71			
DT	01-MAY-2000 (TRENBLrel. 13, Created)			
DT	01-OCT-2000 (TRENBLrel. 15, Last sequence update)			
DT	01-MAR-2003 (TRENBLrel. 23, Last annotation update)			
DE	Pmp 3.			
GN	PMP_3_2			
OS	Chlamydia pneumoniae (Chlamydophila pneumoniae).			
OC	Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.			
OX	NCBI_TaxID=83558;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-J138;			
RX	MEDLINE=20330349; PubMed=10871362;			
RA	Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K., Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;			
RT	"Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CWL029 from USA."			
RL	Nucleic Acids Res. 28:2311-2314(2000).			
DR	EMBL; AP002545; BAA98226.1; -			
DR	InterPro; IPR006315; Autotransport.			
DR	InterPro; IPR005546; Autotransporter.			
DR	InterPro; IPR003368; Chlamydia_PMP.			

P71135 chlamydophi
P77792 chlamydia p
P71132 chlamydophi
P71133 chlamydophi
Q8vu49 chlamydia p
Q8vl57 chlamydia p
Q8vu50 chlamydia p
Q8vu48 chlamydia p
Q9rb67 chlamydia p
P71134 chlamydophi
P71132 chlamydia p
Q9jx66 chlamydia p
Q9rb68 chlamydia p
Q9rb70 chlamydia p
Q9jx68 chlamydia t
Q8vlh7 chlamydia p
Q8vlf9 chlamydia p
Q9jx67 chlamydia p
Q9rb69 chlamydia p
Q98jh8 rhizobium l
Q8tge1 saccharomyc
Q8g9x9 escherichia
Q8g9f46 zymomonas m
Q8vg99 staphylococ
Q8ifx6 caenorhabdi
Q8uau1 agrobacteri
Q5164 saccharomyc
Q25063 helicobacte
Q76602 caenorhabdi
Q9rn12 haemophilus
Q8ydm7 bruceella me
Q9881 chlamydia p
Q8pkm0 xanthomonas
Q8fusi1 bruceella su
Q8y366 ralstonia s
Q8p942 xanthomonas
Q98kc9 rhizobium l
Q8tfg9 rhizosacch
Q8e20 rhizobium l
Q9rb61 chlamydia p
Q8nuj3 staphylococ
Q99y44 staphylococ
Q8zcs9 yersinia pe
Q9f292 yersinia pe

DR Pfam: PF03797; Autotransporter; 1.
 DR Pfam: PF02415; DUF145; 1.
 DR TIGRFAMs: TIGR01414; autotrans_bar1; 1.
 DR TIGRFAMs: TIGR01376; POMP_repeat; 6.
 SQ SEQUENCE 772 AA; 82931 MW; 484FC56D35801EB CRC64;

Alignment Scores:
 Pred. No.: 2,73e-226 Length: 772
 Score: 3991.00 Matches: 771
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.87% Mismatches: 0
 Query Match: 73.78% Indels: 0
 DB: 16 Gaps: 0

US-09-428-122-1 (1-3000) x Q9RB71 (1-772)

QY	569	TTGCTCTTCAGAAAAAATTTTCAACGGATATGCGGTGCTTACCGCAAAACTCTT	628
Db	1	MetLeuPheSerLysAsnPheSerThrAspAsnGlyGlyAlaIleThrAlaLysThrLeu	20
QY	629	TCATTAAACAGGACTACAAATGTCAGCTCTGTTTCTGAAATACCTCTCAAGAAAGGC	688
Db	21	SerLeuThrGlyThrThrMetSerAlaLeuPheSerGluAsnThrSerSerLysLysGly	40
QY	689	GGAGCCATTACAGACTCCGATCCCTTACCATTACCTGGAACCAAGGGGAAGTCTCTTT	748
Db	41	GlyAlaIleGlnThrSerAspAlaLeuThrIleThrGlyAsnGlnGlyValSerPhe	60
QY	749	TCAGCAATACCTCTTCGAGTTCGGAGCTGCAATTTTACAGAGCCCTCGGTGACTATT	808
Db	61	SerAspAsnThrSerSerAspSerGlyAlaAlaIlePheThrGluAlaSerValThrIle	80
QY	809	TCATAAATGCTAAAGTTCCTTTATTGACAAATAGGTACAGGAGCGAGCTCTCAACA	868
Db	81	SerAsnAsnAlaLysValSerPheIleAspAsnLysValThrGlyAlaSerSerThr	100
QY	869	ACGGGGGATATGTCAGGAGGTGCTATCTGTGCTTTATAAACTAGTACAGATCAAGTTC	928
Db	101	ThrGlyAspMetSerGlyGlyAlaIleCysAlaIleLysThrSerThrAspThrLysVal	120
QY	929	ACCTCAGTGGAAATCAGATGTTACTCTTCAGCAACAATACATCGACACACGGGAGGA	988
Db	121	ThrLeuThrGlyAsnGlnMetLeuLeuPheSerAsnAsnThrSerThrAlaGlyGly	140
QY	989	GCTATCTATGTGAAAGCTCGAAGTGGCTTCGGAGGACTTACCCCTATTACAGTAAAT	1048
Db	141	AlaIleThrValLysLysLeuGluLeuAlaSerGlyGlyLeuThrLeuPheSerArgAsn	160
QY	1049	AGTGTCAATGGAGGTACAGCTCCTAAAGGTGGAGCCATAGCTATCGAAGATAGTGGGAA	1108
Db	161	SerValAsnGlyGlyThrAlaProLysGlyGlyAlaIleAlaIleGluAspSerGlyGlu	180
QY	1109	TTGAGTTTATCCCGCATAGTGGTACATGCTCTTTTGGGAATACAGTCACTTCTACT	1168
Db	181	LeuSerLeuSerAlaAspSerGlyAspIleValPheLeuGlyAsnThrValThrSerThr	200
QY	1169	ACTCCTGGGACGAATAGAACTAGTATCGACTTAGGAACAGAGTCAAGATGACAGCTTTG	1228
Db	201	ThrProGlyThrAsnArgSerSerIleAspLeuGlyThrSerAlaLysMetThrAlaLeu	220
QY	1229	CGTTCTGCTCGGTAGAGCCATCTACTTCTATGATCCCATTAACACTACAGATCATCACA	1288
Db	221	ArgSerAlaAlaGlyArgAlaIleThrPheThrAspProIleThrThrGlySerSerThr	240
QY	1289	ACAGTTACAGATGCTTTAAAGTTAATGACTCCCGCAGATTCGCTACATATACA	1348
Db	241	ThrValThrAspValLeuLysValAsnGluThrProAlaAspSerAlaLeuGlnThr	260
QY	1349	GGGAACATCATCTTCACAGGAGAAAGTTATCAGACAGAGCGCGCAGATTTCTAAAAAT	1408
Db	261	GlyAsnIleIlePheThrGlyGlyLysLeuSerGluThrGluAlaAlaAspSerLysAsn	280
QY	1409	CTTACTTCGAAGCTACTACAGCCCTGTAACTCTTTTCAGGAGGTACTCTATTCTTAAACAT	1468

Db	281	LeuThrSerLysLeuLeuGlnProValThrLeuSerGlyGlyThrLeuSerLeuLysHis	300
QY	1469	GGAGTGAATCTGCAGACTCAGGCATTCACACAGCAGACATTCCTGCTCGAAATCGAC	1528
Db	301	GlyValThrLeuGlnThrGlnAlaPheThrGlnGlnAlaAspSerArgLeuGluMetAsp	320
QY	1529	GTAGGAATCTACTAGAACCTGTGATACCTAGACACCAATAACAAATTTGGTCATTAAATC	1588
Db	321	ValGlyThrThrLeuGluProAlaAspThrSerThrIleAsnAsnLeuValIleAsnIle	340
QY	1589	AGTTCTATAGCGGTGCAAGAGGCAAAATAGAAACCAAGCTAGCTCAAAAATCTG	1648
Db	341	SerSerIleAspGlyAlaLysLysAlaLysIleGluThrLysAlaThrSerLysAsnLeu	360
QY	1649	ACTTTATCTGGAACCATCCTATTATGGACCCGACGGCGACGTTTATGAAATCATAGT	1708
Db	361	ThrLeuSerGlyThrIleThrLeuLeuAspProThrGlyThrPheTyrGluAsnHisSer	380
QY	1709	TTAAGAAATCTCAGTCTCCTAGCAGCATCTTAGAGCTCAAGCTTCTGGAACCTGTAACAGC	1768
Db	381	LeuArgAsnProGlnSerTyrAspIleLeuGluLeuLysAlaSerGlyThrValThrSer	400
QY	1769	ACCGCAGTGACTCCAGATCCTATAATGGGTGAGAAATTCATATTACGGCTATCAGGAACT	1828
Db	401	ThrAlaValThrProAspProIleMetGlyGluLysPheHisTyrGlyTyrGlnGlyThr	420
QY	1829	TGGGGCCCAATGTTGGGGACAGGGGCTTCTACGACTGCAACCTTCAACTGACACTAAA	1888
Db	421	TrpGlyProIleValTrpGlyThrGlyAlaSerThrThrAlaThrPheAsnTyrThrLys	440
QY	1889	ACTGGCTATATCTTAATCCCGAGGATCCGCTCTTTAGTCCCTTAATAGCTTATGGAT	1948
Db	441	ThrGlyTyrIleProAsnProGluArgIleGlySerLeuValProAsnSerLeuTrpAsn	460
QY	1949	GCATTTATAGATATTAGCTCTCTCCATTATCTTAGGAGACTGCAAAACCAAGGGTTGCAG	2008
Db	461	AlaPheIleAspIleSerSerLeuHisTyrLeuMetGluThrAlaAsnGluGlyLeuGln	480
QY	2009	GGAGACCTGTCTTTTGGTGTCTGGATTCTAACTTCTTCCATPAAGATAGTACAAA	2068
Db	481	GlyAspArgAlaPheTrpCysAlaGlyLeuSerAsnPhePheHisLysAspSerThrLys	500
QY	2069	ACACACGGGGTTTCGCCATTTAGTGGCGGTATGTCATAGGAGGAAACCTACATCT	2128
Db	501	ThrArgArgGlyPheArgHisLeuSerGlyGlyTyrValIleGlyGlyAsnLeuHisThr	520
QY	2129	TGTTCCAGATAGATTTAGTCTCATTTTGTCTCAGCTCTTTTGGAGAGATAGAGACTAC	2188
Db	521	CysSerAspLysIleLeuSerAlaAlaPheCysGlnLeuPheGlyArgAspArgAspTyr	540
QY	2189	TTTGTAGCTAAGAAATCAAGGTACAGTCTACGGAGAACTCTCTATTACAGACACAGAA	2248
Db	541	PheValAlaLysAsnGlnGlyThrValTyrGlyGlyThrLeuTyrTyrGlnHisAsnGlu	560
QY	2249	ACCTATATCTCTTCTTCCTGCAAACTACGGCTTCTGCTGCTCTTATGTTCTCCTACAG	2308
Db	561	ThrTyrIleSerLeuProCysLysLeuArgProCysSerLeuSerTyrValProThrGlu	580
QY	2309	ATTCTCTGTCTCTTTTCAGAAACCTTAGCTACCCATACCGATAACCATCTGAAAAAC	2368
Db	581	IleProValLeuPheSerGlyAsnLeuSerTyrThrHisThrAspAsnAspLeuLysThr	600
QY	2369	AXGATATACAACTATCTACTCTTAAAGGAAGCTGGGGAATGATAGTTTCCTTTAGAA	2428
Db	601	LysTyrThrThrTyrProThrValLysGlySerTrpGlyAsnAspSerPheAlaLeuGlu	620
QY	2429	TTCCGTGGAGAGCTCCGATTTTGTCTTAGATGAAAGTCTCTATTTCAGCAGCTACATGCC	2488
Db	621	PheGlyGlyArgAlaProIleCysLeuAspGluSerAlaLeuPheGluGlnTyrMetPro	640
QY	2489	TTTCATGAATTCAGTTTCTCTATGCACATCAGGAGGTTTAAAGAACAGGGAACAGAA	2548

```

Db      641  PheMetLysLeuGlnPheValTyAlaHisGlnGluGlyPheLysGluGlnGlyThrGlu 660
Qy      2549  GCTGGTGAATTTGGAAGTAGCGCTTGTGTAATCTTGCTTACCTATCGGATCCGATTT 2608
Db      661  AlaArgGluPheGlySerArgLeuValAsnLeuAlaLeuProIleGlyLeArgPhe 680
Qy      2609  GATAAGGATCAGACTGCCAGAGTCAAGTACATCTAATCTTGTTATCTGCTGAT 2668
Db      681  AspLysGluSerAspCysGlnAspAlaThrTyrAsnLeuThrLeuGlyTyrThrValAsp 700
Qy      2669  CTTGTTCTAGTAAACCCGACTGTACGACACACTGCAATTTAGCGGTGATCTTGAAA 2728
Db      701  LeuValArgSerAsnProAspCysThrThrThrLeuArgIleSerGlyAspSerTrpLys 720
Qy      2729  ACCTTTCGGTAGCAATTTGGCAAGCAAGCTTTAGTCTTCGTGCGAGGAACATTTTGC 2788
Db      721  ThrPheGlyThrAsnLeuAlaArgGlnAlaLeuValLeuArgAlaGlyAsnHisPheCys 740
Qy      2789  TTTAACTCAAAATTTGAAGCCTTTAGCCAAATTTCTTTTGAATTCGGTGGTCACTCGC 2848
Db      741  PheAsnSerAsnPheGluAlaPheSerGlnPheSerPheGluLeuArgGlySerSerArg 760
Qy      2849  AATTACAATGAGCTTAGGAGCAAAATACCAATTC 2884
Db      761  AsnTyrAsnValAspLeuGlyAlaLysTyrGlnPhe 772

RESULT 2
P71135
ID      P71135      PRELIMINARY;      PRT;      926 AA.
AC
DT      01-FEB-1997 (TrEMBLrel. 02, Created)
DT      01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Putative 98 kDa outer membrane protein.
OS      Chlamydomophila abortus.
OC      Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OX      NCBI_TaxID=83555;
RN
RP      SEQUENCE FROM N.A.
RC      STRAIN=ovine abortion S26/3;
RA      Longbottom D., Russell M., Dunbar S.M., Jones G.E., Herring A.J.;
RT      "98kDa protein genes from ovine abortion strain S26/3 Chlamydia
RT      psittaci.";
RL      Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
DR      EMBL; J72499; AB18188.1; -.
DR      InterPro; IPR006315; Autotransporter.
DR      InterPro; IPR005546; Autotransporter.
DR      InterPro; IPR003368; Chlamydia_PMP.
DR      Pfam; PF03797; Autotransporter_1.
DR      Pfam; PF02415; DUF145; 2.
DR      TIGRFAMs; TIGR01414; autotrans_bar1; 1.
DR      TIGRFAMs; TIGR01376; POMP repeat; 6.
SQ      SEQUENCE 926 AA; 98439 MW; 3E755E2F594750F CRC64;

Alignment Scores:
Pred. No.:      1.86e-104      Length:      926
Score:      1915.00      Matches:      410
Percent Similarity:      59.27%      Conservative:      156
Best Local Similarity:      42.93%      Mismatches:      333
Query Match:      35.40%      Indels:      56
Db:      2      Gaps:      18

US-09-428-122-1 (1-3000) x P71135 (1-926)
Qy      101  ATGAGTCTCTTCCCAAGTTGTGATTTTCTACATTTGCTATTTTCCCTTTGTTCTATG 160
Db      1  MetArgProSerLeuTyrLysIleLeuLeuSerSerThrLeuThrLeuProIleSerPhe 20
Qy      161  -----ATTGCTACCGAGACAGACTTTTGGATTCAACT 190
Db      21  HisPheSerGlnLeuHisAlaGluValAlaLeuThrGlnGluSerIleLeuAspAla--- 39
Qy      191  GCGAGTTTCGATGGGAATAAAATGTTAATTTTTCAGTTCGTGAGAGTCAAGAGATGCT 250

```

```

Db      40  -----AsnGlyAlaPheSerProGlnSerThrSerThrAlaGly 52
Qy      251  GGAACCTACCTACCTATTAAAGGAAATGTCACTCTAGAAAATATTCTCTGGAACAGGCACA 310
Db      53  GlyThrIleTyrAsnValGluSerAspIleSerIleValAspVal---GlyGlnThrAla 71
Qy      311  GCAATCAAAAAGCTGTTTAAACAACACTAAGGGCGATTTGACTTTTCACAGGTAAACGGG 370
Db      72  AlaLeuAlaSerSerAlaPheValGlnThrAlaAspAsnLeuThrPheLysGlyAsnAsn 91
Qy      371  AACTCTCTATTGTTCCAAACCGTGGATGCGAGGACTGTAGCAGGGGTCTGTTTAAACAGC 430
Db      92  HisSerLeuSerIleThrAsnAlaAsnAlaGly---AlaAsnProAlaGlyIleAsnVal 110
Qy      431  AGCGTGGTAGATAAATCTACCAACGTTTTATAGGGTTTTCTTCGCTATCTTTTATTCGGTCT 490
Db      111  AsnThrAlaAspLysIleLeuThrLeuThrAspPheSerLysLeuSerPheLysGluCys 130
Qy      491  CCTGGAAGTTCGATAACTACCGGCAAGAGCCCTTAGCTCTCTACGGGTAGCTTGAGT 550
Db      131  ProSerSerLeuValAsnThrGlyLysGlyAlaMetLys---SerGlyGlyAlaLeuAsn 149
Qy      551  TTGACAAAATGTCAGTTTCTCTCAGCAAAAATCTTTCAACGGATAATGGCGGTGCT 610
Db      150  LeuAlaAsnAsnAlaSerIleLeuPheAspGlnAsnTyrSerAlaGluAsnGlyGlyAla 169
Qy      611  ATCACCGCAAAAATCTTTTCAATTAACAGGGGACTACAATGTGCTGCTGTTTCTGAAAAT 670
Db      170  IleSerCysLysAlaPheSerLeuThrGlySerSerLysGluIleSerPheThrThrAsn 189
Qy      671  ACCTCTCTCAAGAAGGGGAGCCATTACAGACTTCCGATGCCCTTACCATTTACTGGAAC 730
Db      190  SerThrAlaLysLysGlyGlyAlaIleAlaAlaThrGlyIleAlaHisLeuSerAspAsn 209
Qy      731  CAAGGGGAAGTCTCTTTTCTGACAATACTTCTTCGGATTCTCGAGCTGCAATTTTACA 790
Db      210  GlnGlyThrIleArgPheSerGlyAsnThrAlaValAsnSerGlyGlyAlaValTyrSer 229
Qy      791  GAAGCTCGGTGACTATTCTTAATAATGCTAAAGTTCCTTTATTTGACAATAAGTCAACA 850
Db      230  GluAlaSerMetThrIleAlaGlyAsnAsnHisValAlaPheSerAsnAsnAlaValSer 249
Qy      851  GGAGGGAGCTCTCAACAACGGGGGATATGTCAGGAGGTGCTATCTGCTTATTAAGT 910
Db      250  GlySerSer-----AspGlyCysGlyGlyAlaIleHisCysSerLysThr 264
Qy      911  AGTACAGATATAAGTCACTGGAATCAGATGTTACTCTTCAGCAACAATAACA 970
Db      265  GlySerAlaProThrLeuThrIleArgAspAsnLysValLeuIlePheGluGluAsnThr 284
Qy      971  TCGACAACAGCGGAGGAGCTATCTATGTGAAAAGCTCGAAGTGGCTTCGGAGAGATT 1030
Db      285  SerSerAlaLysGlyGlyAlaIleTyrThrAspLysLeuLeuLeuThrSerGlyGlyPro 304
Qy      1031  ACCCTATTTCAGTAGAATAGTGTCAATCGAGGTACAGCTCTCAAGTGGGCGCATAGCT 1090
Db      305  ThrAlaPheIleAsnAsnLysValThrHisAlaThr---ProLysGlyGlyAlaIleGly 323
Qy      1091  ATCGAAGATAGTGGGGAATTGAGTTTATCCCGCGAGTAGGTGGTGCATTTGCTTTTAGGG 1150
Db      324  IleAlaAlaAsnGlyGluCysSerLeuThrAlaGluHisGlyAspIleThrPheAspAsn 343
Qy      1151  AATACAGTCACTCTTACTACTCTCTGGGAGC---AATAGAGTAGTATCATCTTAGGAACG 1207
Db      344  AsnLeuMetAlaThrGlnAspAsnAlaThrIleLysArgAsnAlaIleAsnIleGluGly 363
Qy      1208  AGTCAAGATGACAGCTTTGGTCTCTGCTGTAGAGCCATCTACTTCTATGATCCC 1267
Db      364  AsnGlyLysPheValAsnLeuArgAlaAlaSerGlyLysThrIleSerPheTyrAspPro 383
Qy      1268  ATAACCTACAGGATCATCCACAACAGTTACAGATGCTTAAAGATTAAATGAGACTCCGCA 1327

```

Db 384 IleThrValGluGlyAsn-----AlaAlaAspLeuLeuThrLeuAsnLysAlaGluGly 401
 QY 1328 GATTCTGCACATAATACAGGGAACATCATCTTCACAGGAGAGAAAGTATTCACAGACA 1387
 Db 402 AspLysThr-----TyrAsnGlyArgIleIlePheSerGlyGluLysLeuThrGluGlu 419
 QY 1388 GAGGCGCGAGATCTTAAATCTTACTCGAAGCTACTACAGCCGTGAATCTTCAGGA 1447
 Db 420 GlnAlaAlaValAlaAspAsnLeuLysThrThrPheThrGlnProIleThrLeuAlaAla 439
 QY 1448 GGTACTCTATCTTAAACATCGAGTGCATCTGCAGACTCAGGCATTCATCAACAGGCA 1507
 Db 440 GlyGluLeuValLeuArgSerGlyValGluValGluAlaLysThrValValGlnThrAla 459
 QY 1508 GATTCTCGTCTGAAATGACGTAGGAATCTACTAGAA---CCTGCTCATCTAGCAC 1564
 Db 460 GlySerLeuLeuLeuMetAspAlaGlyThrLysLeuSerAlaLysThrGluAspAlaThr 479
 QY 1565 ATAACAATTTGGTCTATTAACTAGTCTATAGACGGTGAAGAGCAAAATAGAA 1624
 Db 480 LeuThrAsnLeuAlaIleAsnProAsnThrLeuAspGlyLysLysPheAlaValValAsp 499
 QY 1625 ACCAAGCTACGTCAAAATCTGACTTTATCTGGAACCATCATCTTATGGACCGGAG 1684
 Db 500 AlaValAlaAlaGlyLysAsnValThrLeuSerGlyAlaIleGlyValIleAspProThr 519
 QY 1685 GGCACGTTTATGAAATCATAGTTTAAAGAAATCTCTAGTCTAGCACATCTTAGAGCTC 1744
 Db 520 GlyLysPheTyrGluAsnHisLysLeuAsnAspThrLeuAlaLeuGlyIleGlnLeu 539
 QY 1745 AAGCTTCTGGAACTGTATACAGCAGCGAGTGCATCTAGATCCTTAATAGGT---GAG 1801
 Db 540 SerGlyLysGlySerValThrThrAsnVal---ProSerHisValValGlyValAla 558
 QY 1802 AAATTCATCTACGCTATCAGGAACTTGGGCCCAATTTGTTGG-----1846
 Db 559 GluThrHisTyrGlyTyrGlnGlyAsnTrpSer---ValSerTrpValLysAspAsnAsn 577
 QY 1847 GGGACAGGGGCTTCTACGACTGCAACCTTCAACTGAGCTAAACTGGCTATATTCCTAAT 1906
 Db 578 SerAspProLysThrGlnThrAlaIlePheThrTipAsnLysThrGlyTyrValProAsn 597
 QY 1907 CCGAGCGTATCGGCTCTTTAGTCCCTATAGCTTATGGAATCATTATAGATATAGC 1966
 Db 598 ProGluArgAlaProLeuValLeuAsnSerLeuTrpGlySerPheIleAspLeuArg 617
 QY 1967 TCTCTCCATATCTTATGGAGACTGCAACGGAAGGG---TTGCAGGAGACCTGCTTTT 2023
 Db 618 SerIleGlnAspValLeuGluArgSerValAspSerIleLeuGluThrArgArgGlyLeu 637
 QY 2024 TGGTGTGCTGGATTATCTACTTCTCCATAGGATAGTACAAACACAGACCGGGTTT 2083
 Db 638 TrpValSerGlyIleGlyAsnPhePheHisLysAspArgAsnAlaGluAsnArgLysPhe 657
 QY 2084 CGCATTTAGTGGCGGTATGTCATAGAGAGAACTCATATCTGTTTCAGATAGATT 2143
 Db 658 ArgHisIleSerSerGlyTyrValLeuGlyAlaThrThrAsnThrSerArgGluAspSer 677
 QY 2144 CTTAGTGCTGCATTTTGTGAGCTCTTTGGAGAGATAGAGACTTCTTGTAGCTAAGAAT 2203
 Db 678 LeuSerValAlaPheCysGlnLeuPheAlaLysAspLysAspTyrLeuValSerLysAsn 697
 QY 2204 CAAGGTACAGTCTACGGAGAACTCTCTATTACCGACACAGCAAACTTATCTCTCT 2263
 Db 698 AlaAlaAsnValTyrAlaGlySerValTyrTyrGlnHisValSerLysPheAspAspLeu 717
 QY 2264 -----CCTTGCAACTACGGCTTGTGTTGTTCTTATGTT 2299
 Db 718 ThrArgLeuPheAsnGlyProAsnThrCys-----CysSerGlyPhe 731
 QY 2300 CCTACAGATTCCTGTTCTCTTTTCAGGAACTTATAGCTACACCCATACGATACGAT 2359
 Db 732 SerLysGluIleProIlePheLeuAspAlaGlnIleThrTyrCysHisThrAlaAsnAsn 751

QY 2360 CTGAAACCAAGTATPACAACTATCTCTACTGTTTAAAGGAAGCTGGGGAATGATAGTTTC 2419
 Db 752 MetThrThrSerTyrThrAspTyrProGluValLysGlySerTrpGlyAsnAspThrLeu 771
 QY 2420 GCTTTAGAAATTCGGTGGAGAGCTCCGATTTGCTTAGATGAAGTCTCTATTTCAGCAG 2479
 Db 772 GlyLeuThrLeuSerThrSerValProIleProValPheSerSerSerIlePheAspSer 791
 QY 2480 TACATGCTCCTTCATGAAATTCAGTTTCTATATGCACATCAGGAAGGTTTAAAGAACAG 2539
 Db 792 TyrAlaProPheAlaLysLeuGlnValValTyrAlaHisGlnAspAspPheLysGluPro 811
 QY 2540 GGAACAGAGCTCGGAATTTGGAAGTAGCCGCTCTGTGAATCTTTCCTTACCTACCGG 2599
 Db 812 ThrThrGluGlyArgValPheGluSerSerAspLeuLeuAsnValSerValProIleGly 831
 QY 2600 ATCCGATTTGATAAGAAATCAGACTGCCAAGATGCAAGTCAATCTCACTCTCTGTTAT 2659
 Db 832 IleLysPheGluLysLeuSerTyrGlyGluArgSerAlaTyrAspLeuThrLeuMetTyr 851
 QY 2660 ACTGGGATCTTGTTCGTAGTAACCCGACTGTACGACACACTCGGAATTAGCGGTGAT 2719
 Db 852 IleProAspValTyrArgHisAsnProSerCysMetThrGlyLeuAlaIleAsnAspVal 871
 QY 2720 TCTTGGAAAACTTCGCTACGAAATTTGGCAAGACAGCTTTAGTCTCTGTCGACGGAAC 2779
 Db 872 SerTrpLeuThrThrAlaThrAsnLeuAlaArgGlnAlaPheIleValArgAlaGlyAsn 891
 QY 2780 CATTTTGTCTTAACTCAAAATTTCAAGCTTTTACCAATTTCTTTGAAATTCGCTGGG 2839
 Db 892 HisIleAlaLeuThrSerGlyValGluMetPheSerGlnPheGlyPheGluLeuArgSer 911
 QY 2840 TCATCTCGCAATPACAAATGTAGACTTAGCAGCAAAAATACCAATTC 2884
 Db 912 SerSerArgAsnTyrAsnValAspLeuGlyAlaLysValAlaPhe 926

RESULT 3

P77792 ID P77792 PRELIMINARY; PRT; 839 AA.
 AC 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE POMP90B precursor.
 OS Chlamydia psittaci (Chlamydia psittaci).
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=83554;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OVINE ENZOOTIC ABORTION ISOLATE S26/3;
 RA Longbottom D., Russell M., Jones G.E., Lainson P.A., Herring A.J.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OVINE ENZOOTIC ABORTION ISOLATE S26/3;
 RA Longbottom D., Russell M., Jones G.E., Lainson P.A., Herring A.J.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U65943; AAC15924.1; -
 DR EMBL; U65942; AAC15922.1; -
 DR InterPro; IPR006315; Autotransport.
 DR InterPro; IPR003368; Chlamydia_PMP.
 DR Pfam; PF02415; DUF145; 1.
 DR TIGRFAMs; TIGR01414; autotrans_bar1; 1.
 DR TIGRFAMs; TIGR01376; POMP_repeat; 3.
 KW Signal.
 FT SIGNAL 1 16 POTENTIAL.
 FT CHAIN 17 839 POMP90B.
 SQ SEQUENCE 839 AA; 89824 MW; 4581C7CBATFF4C4 CRC64;

Alignment Scores:

Pred. No.: 1 81e-89 Length: 839
 Score: 1659.50 Matches: 384

Percent Similarity:	54.17%	Conservative:	136
Best Local Similarity:	40.00%	Mismatches:	287
Query Match:	30.68%	Indels:	153
DB:	2	Gaps:	25
US-09-428-122-1 (1-3000) x P77792 (1-839)			
Qy	89	ATATATTTTCAAGAGTCTCTTCCCAAGTTGTTATTTTCTACATTTGCTATTTTC	148
Db	5	ValTyrTrpPheLeuIleSerSer-----SerLeuPheAlaSerAsn	18
Qy	149	CTTTGCTATGATT--GCTACCGAGACAGTTTTGGATTCAAGTCGCGAGTTTCGATCGG	205
Db	19	SerLeuSerPheAlaAsnAspAlaGlnThrAlaLeuThrProSerAspSerTyrAsnGly	38
Qy	206	AAT---AAAAATGCTAATTTTTCAGTTCGTGAGAGTCAGGAAGATGCTGGAACTACCTAC	262
Db	39	AsnValThrSerGluGluPheGlnValLysGluThr-----SerSerGlyThrThrTyr	56
Qy	263	CTATTTAAGGGAAATGTCACTCTAGAAAATATTCCTGGAACAGCAGCAATCAAAAA	322
Db	57	ThrCysGluGlnValCysIle---SerPheAlaGlyLysAspSerGlyLeuLysLys	75
Qy	323	AGCTGTTTTAACACACTAAGCGGATTTGACCTTTTCACAGGTAAACGGGAACCTCTCTATG	382
Db	76	SerCysPheSerAlaThr---AspAsnLeuThrPheLeuGlyAsnGlyTyrThrLeuCys	94
Qy	383	TTCCAAACGGTGGATGACGAGCTGTAGCAGGGGCTGCTGTTTAAACAGCAGCGTGTAGAT	442
Db	95	PheAspAsnIleThrThrAlaSerAsnProGlyAlaIleAsnValGlnGlyGlnGly	114
Qy	443	AAATCTACACAGTTTATAGGGTTTCTTCCTCTATCTTTTATGCGTCTCTCGAAGTTGG	502
Db	115	LysThrLeuGlyIleSerGlyPheSerLeuPheSerCysAlaTyrCysProGly---	133
Qy	503	ATACTACCGGCAAGAGCGGTAGCTGCTACGGGTAGCTTGATTTGACAAAAT	562
Db	134	--ThrThrGlyTyrGlyAlaIleGln---ThrLysGlyAsnThrThrLeuLysAspAsn	151
Qy	563	GTCAGTTGCTCTTCAGCAAAAATTTTCAACGGTAATGCGGTGCTATCAGCGCAAAA	622
Db	152	SerSerLeuValPheHisLysAsnCysSerThrAlaGluGlyAlaIleGlnCys---	170
Qy	623	ACTCTTTTCATTAAACAGGACTACAAATGTCAGCTCTGTTTCTCAAAATACCTCTCAAG	682
Db	170	-----	170
Qy	683	AAAGCGGAGCCATTACAGACTTCGATGCGCTTACCATTTACTGGAAACCAAGGGGAAGTC	742
Db	170	-----	170
Qy	743	TCCTTTTCTGCAATACTCTCTCGATTCTGGAGCTGCAATTTTACAGAGCCTCGGNG	802
Db	170	-----	170
Qy	803	ACTATTTCTTAATGCTAAAGTTTCTCTTATTGACAATAAGGTCACAGGAGCGAGCTCC	862
Db	170	-----	170
Qy	863	TCAACACGGGGATATGTCAGAGGTGCTATCTGCTTTATAAACTAGTACAGATCT	922
Db	171	-----LysGlySerSerAspAla	176
Qy	923	AAGTCAACCTTCATGGAATTCAGATGTTACTCTTACAGCAACATACATCGACACAGCG	982
Db	177	GluLeuLysIleGluAsnAsnGlnAsnLeuValPheSerGluAsnSerSerThrSerLys	196
Qy	983	GGAGGAGTACTATGTGAAAAGCTCGAACTGCTTCGCGAGACTTACCTCTATTTCAGT	1042
Db	197	GlyGlyAlaIleTyrAlaAspLysLeuThrIleValSerGlyGlyProThrLeuPheSer	216
Qy	1043	AGAAATAGTGTCAATGGAGGTACAGCTCCTTAAAGTGGAGCCATGCTATCGAAGAT---	1099

QY	2138	AAGATTCTTAGTGGCTGCTATTTTCCTCAGCTCTTTGGGAAGAGATAGACACTACTTTGTAGCT	2197
Dd	584	AspilepheSerAlalaPheCyGlnLeuPheGlyAspTyrlleuValSer	603
QY	2198	AGAATAAAGGTACAGTCTACGGAGGAACCTCTATTACCAGCACAAACCACTATATC	2257
Dd	604	LysAsnAlaAsnilefyrAlaGlySerLeutyrryGlnHis-----lle	619
QY	2258	TCTCTTCCTTGCAAACCTACCGCCCTTGTTCGTTCTATATGTTCTCAGAGATTCCTGTT	2317
Dd	620	SerTyrrTrpSerAlaTrpGlnAsnLeuLeuGlnAsnThrIleGlyAlaGluAlaProLeu	639
QY	2318	CTCTTTTCAGCAAACTTACTACCATCACCGATAAACCATGCATCTGAAAAACC-----	2368
Dd	640	ValLeuAsnAlaGlnLeuThryrCysHisAlaSerAsnAspMetLysThrAsnMetThr	659
QY	2369	-----AAGTATACAACATATCTCTTAAGGAAGAGCTGGGGGATGAT	2413
Dd	660	ThrThrTyrrAlaProArgLysThrThrTyrrAlaGluIleLysGlyAspTrpGlyAsnAsp	679
QY	2414	AGTTTCGTTTGAATTCGGTGGAGAGCTCCGATTTGCTTAGATGAAGTCTCTATTT	2473
Dd	680	CysPheGlyValGluLeuGlyAlaThrValProIleGlnThrGluSerSerLeuLeuPhe	699
QY	2474	GAGCAGTACATGCCCTTCATGAATTCGAGTTTGTCTATGCATCAGGAAGGTTTTAAA	2533
Dd	700	AspMetTyrrSerProPheLeuLysPheGlnLeuValHisThrHisGlnAspAspPheLys	719
QY	2534	GAACAGGAACA---GAAGCTCGTGAATTTGGAAGTAGACCGCTCTTGTAATCTTGCCTTA	2590
Dd	720	GluAsnAsnSerAspGlnGlyArgTyrrPheGluSerSerAsnLeuThrAsnLeuSerLeu	739
QY	2591	CCATACGGATCCGATTTGATGAAGATCAGACTCCAGCTGCCAAGATGCAACGTACAATCTAACT	2650
Dd	740	ProIleGlyIleLysPheGluArgPheAlaAsnAsnAspThrAlaSerTyrrHisValThr	759
QY	2651	CTTGTTTATCTGTGGATCTTTGTCGTAGTAACCCCGACTGTACGACAACTCGCAATT	2710
Dd	760	AlaAlaTyrrSerProAspIleValArgSerAsnProaspCysThrThrSerLeuLeuVal	779
QY	2711	AGCGGTGATTCT-----TGAAAACTTCGGTACGAATTTGGCAAGCAAGCTTTAGTC	2764
Dd	780	SerProaspSerAlaValTrpValThrLysAlaAsnAsnLeuAlaArgSerAlaPheMet	799
QY	2765	CTTCGTCGAGGACCATTTTGTCTTAACCTCAATTTTGAGCCTTTAGCCAATTTCT	2824
Dd	800	LeuGlnAlaGlyAsnTyrrLeuSerLeuSerHisAsnIleGluIlePheSerGlnPheGly	819
QY	2825	TTTGAATTCGCTGGGTATCTCGCAATTACAAATGTAGACTTTAGGAGCAAAATACCAATTC	2884
Dd	820	PheGluLeuArgGlySerSerArgThrTyrrAsnValasPheGlySerLysIleGlnPhe	839
 RESULT 4			
ID	P71132	PRELIMINARY;	PRT; 847 AA.
AC	P71132;		
DT	01-FEB-1997 (TREMBLrel. 02, Created)		
DT	01-FEB-1997 (TREMBLrel. 02, Last sequence update)		
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)		
DE	POMF91A.		
OS	Chlamydophila abortus.		
OC	Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydochloa.		
OX	NCBI_TaxID=83555;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=S26/3;		
RX	MEDLINE=96406378; PubMed=8810511;		
RA	Longbottom D., Russell M., Jones G.E., Lainson F.A., Herring A.J.;		
RT	"Identification of a multigene family coding for the 90 kDa proteins		
RL	of the ovine abortion subtype of Chlamydia psittaci."		
RF	JEMS Microbiol. Lett. 142:277-281 (1996).		
RN	[2]		

QY 809 TCTAATAATGCTAAAGTTTCCTTTATTGACAATAAGGTCACAGGAGCGAGCTCCTCAACA 868
Db |||||
186 GluAsn----- 187
QY 869 ACGGGGATATGTCAGGAGGTGCTATCTGTGCTTATAAACTAGTACACACTAGTACAGTAC 928
Db ----- 187
QY 929 ACCCTCACTGGAAATFCAGATGTTACTCTTCAGCAA CAATACATACGACAACAGCGGAGGA 988
Db |||||
188 -----AsnGlnAsnLeuValPheAlaGluAsnSerSerSerSerSerGlyGly 203
QY 989 GCTATCTATGCAAAAAGCTCGAAGCTGCTCCGAGGAGGACTTACCTATTTCAGTAGAAT 1048
Db |||||
204 AlalIeTyrAlaAspLeuThrIleValSerGlyGlyProThrLeuPheSerAsnAsn 223
QY 1049 AGTGTCAATGGAGGTACAGCTCCTAAAGGTGGAGCATAGCTATATCGAAGATAGT---GGG 1105
Db |||||
224 SerValSer---AlaSerSerProLysGlyAlaIleCysIleLysAspSerGlyGly 242
QY 1106 GAATTGAGTTATCCCGCATAGTGGTGACATGTCTTTTATAGGAATACAGTACTTCT 1165
Db |||||
243 GluCysSerLeuThrAlaAspLeuGlyAspIleThrPheAspGlyAsnLysIleLys 262
QY 1166 ACT-----ACTCTGGGACGATAGAGTAGTATCGACTTAGGACGAGTCAAG 1216
Db |||||
263 ThrAsnGlyGlySerProThrValThrArgAsnSerIleAspLeuGlySerSerGlyLys 282
QY 1217 ATGACAGCTTTCGTCTGCTGCTAGAGCATCTACTCTATGATCCCACTATACA 1276
Db |||||
283 PheThrLysLeuAsnAlaGlyGluGlyPheGlyIlePhePheTyrAspProIleThrGly 302
QY 1277 GGATCATCCACAACAGTTACAGATGTCTTAAAGTTAATAGACATCCCGCAGATTCGCA 1336
Db |||||
303 GlyGlySer-----AspGluLeuAsnIleAsnLys-----GlnAspThr 315
QY 1337 CTACAAATACAGGNACTATCTCTCACAGAGAAAGTTATTCACAGACAGAGCCGCA 1396
Db |||||
316 ValAspTyrThrGlyLysIleValPheSerGlyGluArgLeuSerAspGluGlyLys 335
QY 1397 GATTCCTAAAATCTTACTTCGAAGCTACTACAGCTGTAACTCTTTCAGGAGTACTCTA 1456
Db |||||
336 ValAlaAlaAsnLeuLysSerAspPheLysGlnProLeuLysIleGlySerGlySerLeu 355
QY 1457 TCTTTAAACATGGAGTACTCTGACAGCTACGACATTCACCAAGCAGAGTCTCGT 1516
Db |||||
356 IleLeuLysAspGlyValThrLeuGluThrLysSerPheThrGlnThrGluGlyAlaThr 375
QY 1517 CTCGAAATGGAGCTAGGAATCTCTAGAACCT-----GCTGATACTAGCACC 1564
Db |||||
376 ValValMetAspLeuGlyThrThrLeuGlnThrProSerSerGlyGlyGluThrIleThr 395
QY 1565 ATAAACAATTTGCTATTAAACATCAGTTCTATAGACGGTGCA-----AAGAAG 1612
Db |||||
396 LeuThrAsnLeuAspIleAsnValAlaSerLeuGlyGlyGlyValAlaProAspPro 415
QY 1613 GCAAAAATAGAAACCAAGCTACGTCAAAAATCTGACTTTTATCGAACCATCTTTA 1672
Db |||||
416 AlaLysValGluAlaThrThrGluSerLysThrValThrIleAsn---AlaValAsnLeu 434
QY 1673 TTGGACCGACGGCGCGTGTATGAAATCATAGTTTAAAGAAATCCTCAGTCTACGAC 1732
Db |||||
435 ValAspAspAsnGlyAsnAlaTyrGluTyrProIleLeuAlaAlaSerGlnProPheThr 454
QY 1733 ATCTTAGAGCTCAAA-----GCTCTGGAATGTAAACAAGCAGCCAGTG----- 1777
Db |||||
455 AlaIleGluValArgSerGlySerSerGlySerIleThrLysProThrThrAsnLeuGlu 474
QY 1778 -----ACTCCAGATCCTATAATGGGTGAGAAATTCATTAGCGCTATCAGGGAATCTGG 1831
Db |||||
475 AsnTyrThrProProThr-----HisTyrGlyTyrGlnGlyAsnTrp 488

QY 1832 GGCCCAATTTGTTGGGACAGGGGCTTCTACG-----ACTGCAACCTTCAACTGG 1882
Db ---ThrValThrTrpLysGlnGlySerSerAlaGlnGluLysThrAlaThrLeuThrTrp 507
QY 1883 ACTAAAATCGCTATATTCTTAATCCCGAGCGTATCGGCTCTTTAGTCCCTTAATAGCTTA 1942
Db |||||
508 GluGlnThrGlyTyrSerProAsnProGluArgGlnGlySerLeuValProAsnThrLeu 527
QY 1943 TGGAAATGCATTTATAGATATTAGTCTCTCTCATTTATCTTATGGAGAGTCGAAACGAAGG 2002
Db |||||
528 TrpGlySerPheSerAspIleArgAlaIleGlnAsnLeuMetAspIleSerValAsnGly 547
QY 2003 TTGCAGGAGAGCGGCTGCTTTTGTGTCTGATTTATCTAACTTCTTCCATAAGGATAGT 2062
Db |||||
548 AlaAspTyrHisArgGlyPheTrpValSerGlyLeuGlyAsnPheLeuHisLysSerGly 567
QY 2063 ACAAAAACACAGACCGGGTTCGCGCATTTAGTGGCGGTATATGTCATAGGAGGAAACCTA 2122
Db |||||
568 SerAspThrLysArgLysPheArgHisAsnSerAlaGlyTyrAlaLeuGlyValTyrAla 587
QY 2123 CATACTTGTTCAGATAGATTTCTTAGTCTGCATTTGTCTAGCTCTTTGGAGAGATAGA 2182
Db |||||
588 GlnThrProSerGluAspValPheSerAlaAlaPheCysGlnLeuPheGlyLysAspLys 607
QY 2183 GACTACTTTGTAGCTAAGAATCAAGGTACAGTCTACGAGGAACTCTCTATTACCAGCAC 2242
Db |||||
608 AspTyrLeuValSerLysAsnSerThrValTyrAlaGlySerIleTyrGlnHis 627
QY 2243 AACGAAACCTATATCTCTTCTTCTTGTGCAACTCGGCGCTTGTCTGTTTATTTCTCT 2302
Db |||||
628 -----GAGATTCTCTCTCTCTTTCAGGA 2329
QY 633 ThrTrpAsnThrLeuLeuGlnAsnThrLeuGlyAlaGluAlaProLeuValLeuAsnAla 652
Db |||||
2330 AACCTTAGCTACCCATACCGATAACGATCTGAAAAACCAAGTAT----- 2374
QY 653 GlnLeuAlaTyrCysHisAlaSerAsnAsnMetLysThrAsnMetThrAspThrTyrAla 672
Db |||||
2375 -----ACAACATATCTCTGTTTAAAGAGAGCTGGGGAAATGATGTTTCCCTTTA 2425
QY 673 ProProLysThrThrTyrSerGluIleLysGlyAspTrpGlyAsnAspCysPheGlyVal 692
Db |||||
2426 GAATTCGGTGAAGAGCTCCGATTTGCTTAGTAAAGTGTCTATTATGACGAGTACATG 2485
QY 693 GluPheGlyAlaLysAlaProIle---GluThrAlaSerLeuLeuPheAspMetTyrSer 711
Db |||||
2486 CCTTCATGAATTCAGTGTGTCTATGCACATCAGGAAGGTTTAAAGAACAGGGAACA 2545
QY 712 PropheValLysLeuGlnLeuValHisAlaHisGlnAspAspPheLysGluAsnAsnSer 731
Db |||||
2546 ---GAAGCTCGTGAATTTGGAAGCGCTCTTGTGAATCTTTCCTTACCTATCCGGATC 2602
QY 732 AspGlnGlyArgTyrPheGluSerAsnAsnLeuThrAsnLeuSerMetProIleGlyVal 751
Db |||||
2603 CGATTTGATAGGAATCAGATCGCAAGTACCACTCACTCACTTCTTCTGTTTACT 2662
QY 752 LysLeuGluLysPheSerHisLysAspThrAlaSerTyrAsnLeuThrLeuAlaTyrAla 771
Db |||||
2663 GTGGATCTTGTTCGTAGTAACCCGACTGTACGACCAACTCGCAATTTAGCGGTGATCT 2722
QY 772 ProAspIleValArgSerAsnProAspCysThrAlaSerLeuLeuValSerProThrSer 791
Db |||||
2723 -----TGGAAAACTTCGTAACGAATTTGGCAGACAAAGCTTTTCTCTGTCAGGG 2776
QY 792 AlaValTyrValThrLysAlaAsnAsnLeuAlaArgHisAlaPheIleLeuGlnAlaGly 811
Db |||||
2777 AACCAATTTTGTCTTAATCTCAAAATTTTGAAGCTTTTAGCCATTTTCTTTGAAATGGCT 2836
QY 812 AsnTyrLeuAlaLeuThrArgAsnThrGluLeuPheSerGlnPheGlyPheGluLeuArg 831
Db |||||
2837 GGGTCATCTCGCAATTACATGTAGACTTAGGAGCAAAATACCAATTC 2884

832 GlySerCysArgThrTyrAsnIleAspLeuGlySerLysIleGlnPhe 847

RESULT 5

P71133 PRELIMINARY; PRT; 846 AA.

AC P71133;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE POMP91B precursor
OS Chlamydomophila abortus.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
NCBI_TaxID=83555;
[1]
RN SEQUENCE FROM N.A.
RP
RC STRAIN=S26/3;
RE MEDLINE=96406378; PubMed=9529048;
RA Longbottom D., Russell M., Jones G.E., Lainson F.A., Herring A.J.;
RT "Identification of a multigene family coding for the 90 kDa proteins
of the ovine abortion subtype of Chlamydia psittaci.";
RL FEMS Microbiol. Lett. 142:277-281(1996).
[2]
RN SEQUENCE FROM N.A.
RP
RC STRAIN=S26/3;
RE MEDLINE=98187897; PubMed=9529048;
RA Longbottom D., Russell M., Dunbar S.M., Jones G.E., Herring A.J.;
RT "Molecular cloning and characterization of the genes coding for the
highly immunogenic cluster of 90-kilodalton envelope proteins from the
Chlamydia psittaci subtype that causes abortion in sheep.";
RL Infect. Immun. 66:1317-1324(1998).
DR EMBL; U65943; AAC15923.1; -;
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF02415; DUF145; 1.
DR TIGRFAMs; TIGR01376; POMP_repeat; 3.
KW SIGNAL.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 846 POMP91B.
SQ SEQUENCE 846 AA; 90834 MW; 4CDC31DC03C2964E CRC64;

Alignment Scores:
Pred. No.: 2,87e-86 Length: 846
Score: 1605.00 Matches: 375
Percent Similarity: 53.02% Conservative: 143
Best Local Similarity: 38.38% Mismatches: 279
Query Match: 29.67% Indels: 180
DB: 2 Gaps: 28

US-09-428-122-1 (1-3000) x P71133 (1-846)

QY 89 ATATATTTTACAAATGAAGTCCTCTTCTCCCAAGTTTGTTATTTCTACATTTCTGCTATTTC 148
Db 5 ValTyrTrpPheLeuIleSerSer-----SerLeuLeuAlaSerAsn 18
QY 149 CCTTTGCTTATGATT--GCTACCGAGACAGATTTTGGATTCAAGTCGCGAGTTTCGATGGG 205
Db 19 SerLeuSerPheAlaGlnValThrAsnGluThrLeuThrSerSerAspSerTyrAsnGly 38
QY 206 AAT--AAAAATGGTAATTTTTCAGTTCGTAGAGTCAGGAAGATGCTGGAACTACCTAC 262
Db 39 AsnValThrSerAspGluPheGluValIleGluThr-----ThrSerGlyAlaIleTyr 56
QY 263 CTATTTAAGGGAATGTCACCTCAGAAATATTTCTGGAAACAGGACAGCAATCACAAA 322
Db 57 ThrCysGluGlyAsnValCysIle---SerTyrAlaGlyLysAspSerProLeuAsnLys 75
QY 323 AGCTGTTTTAAACAACTAAGGCGCATTTTCAGCTTTTCACAGTAACCGGAACCTCTCTATTG 382
Db 76 SerCysPheSerGluThrThrGluAsnLeuSerPheIleGlyAsnGlyTyrThrLeuCys 95
QY 383 TTCCAAACCGGTGGATGCAGGACTGTAGCAGGGGGCTGTGTTAAACAGACAGCTGGTAGAT 442
Db 96 PheAspAsnIleThrThrGlnSerSerHisProGlyValaIleSerValSerGlyThrAsn 115

[illegible]

```
QY 1502 CAGCAGATTCTCGTCTCGAATGAGCTAGGACTACTCTAGAA-----CCT 1549
Db 372 ThrLysGlySerThrValValMetAspLeuGlyThrThrLeuGlnThrProSerSerSer 391
QY 1550 GCTGATACACCATCAACAAATTTGGTCTATTAAACATCAGTCTCTATA-----GAC 1600
Db 392 GlyGluThrThrLeuThrAsnLeuAspLeuAsnLeuLeuGlyGlyGly 411
QY 1601 GGTCAAGAAGGCAAAATAGAACCAAGCTAGCTGCAAAATCTGACTTTATCTGGA 1660
Db 412 GlyThrAlaProAlaLysLeuAlaThrAsnThrAlaSerGlnAlaIleSerIleAla--- 430
QY 1661 ACCATCATTTATGGACCGGACGGCAGGCTTTATGAATATCATAGTTTAAAGAAATCCT 1720
Db 431 AlaValAsnLeuValAsnThrAspSerAsnThrThrGluAspProIleLeuSerAlaSer 450
QY 1721 CAGTCTACGACATCTTTAGAGCTCAAAAGCTTCTGGAAGTGAACACGACGCGAGTGACT 1780
Db 451 LysSerPheSerAlaIle-----ThrAlaThrThrSerSerSerThrValThr 466
QY 1781 CCAGATCCTATAATGGGTGAGAAATTC-----CATTCGGCTCATCAGGAACT 1828
Db 467 ProProGluThrAsnLeuLysAsnThrThrProThrHisThrGlyThrGlnGlyAsn 486
QY 1829 TGGGGCCCAATTTGTTGGGGACAGGGGCTCTACG-----ACTGCAACCTTCAAC 1879
Db 487 Trp---ThrValThrTrpLysGlnLysSerSerAlaGlnGluLysThrAlaThrLeuThr 505
QY 1880 TGGACTAAATCGCTATATCTTAATCCGAGCGTATCGGCTCTTTAGTCCCTAATAGC 1939
Db 506 TrpGluGlnThrGlyThrSerProAsnProGluArgValGlySerLeuValProAsnThr 525
QY 1940 TTATGGAATGATTTATAGATTAATAGCTCTCTCCATTATCTTATGGAAGTCAACACAA 1999
Db 526 LeuTrpGlyAlaPheSerAspThrArgAlaIleGlnAsnLeuMetAspIleSerValAsn 545
QY 2000 GGGTTGCAGGAGACCGGCTTTTGGTGCTGGATTATCTTAATCTCTTCCATAGGAT 2059
Db 546 GlyAlaAspThrSerArgGlyPheTrpValSerSerLeuAlaAsnPheLeuAsnLysSer 565
QY 2060 AGTACAAAAACACGACGGGGTTTCGCCATTTGAGTGGCGGTATGTATGAGGAGAAAC 2119
Db 566 GlySerAspThrLysArgLysPheArgHisSerAlaGlyThrAlaLeuGlyValThr 585
QY 2120 CTACATATTTGTCAGATAAGATTTCTAGTGTGCTGCTATTTGCTGCTTTGGAAGAGAT 2179
Db 586 AlaGlnThrProSerAspValCysSerAlaAlaPheCysGlnLeuPheGlyLysAsp 605
QY 2180 AGAGACTACTTTGAGTAAAGATCAAGTACAGTACAGTACGGAGGAACTCTCTATTACCAG 2239
Db 606 LysAspThrPheValSerLysAsnSerSerThrIleThrAlaGlySerIleThrThrGln 625
QY 2240 CACAACGAAACCTATATCTCTCTCTGCAAACTACGGCCTTGTGTTGTTCTTATGTT 2299
Db 626 His-----IleSerThrTrp 630
QY 2300 CCTACA-----GAGATTCCTGTTCTCTTTCA 2326
Db 631 AsnThrTrpAsnThrLeuLeuGlnAsnThrLeuGlyAlaGluAlaProLeuValLeuAsn 650
QY 2327 GGAACCTTAGCTACACCATCAGGATACGATCTCGAAACCAAGTATACA----- 2377
Db 651 AlaGlnLeuThrThrCysHisAlaSerAsnMetLysThrAsnMetThrAsnThrThr 670
QY 2378 -----ACATATCTACTGTTTAAAGGAAGCTGGGGAATCATAGTTTCGCT 2422
Db 671 ThrProLysAsnValThrProSerGluIleLysGlyAspTrpGlyAsnAspCysPheGly 690
QY 2423 TTAGATTCGGTGAAGACTCCGATTTGCTTTAGTAGAAGTCTCTATTTCAGCAGGTAC 2482
Db 691 ValGluPheGlyAlaLysAlaProIle---GluThrAlaSerLeuLeuPheAspMetThr 709
QY 2483 ATGCCCTTCATGAATTCAGTTTGTCTATGCATCAGGAAGGTTTAAAGACAGGA 2542
```

```
Db 710 SerProPheValLysLeuGlnLeuValHisAlaHisGlnAspAspPheLysGluAsnAsn 729
QY 2543 ACA---GAAGCTCGTGAATTTGGAAGTAGGCGGTCTTGTGAATCTTGCCTTACCTATCGGG 2599
Db 730 SerAspGlnGlyArgThrPheGluSerAsnLeuThrAsnLeuSerMetProIleGly 749
QY 2600 ATCCGATTTGATAGGAATCAGATGCGCAAGATGCAAGCTACCAATCTCTTGGTTAT 2659
Db 750 ValLysLeuGluLysPheSerHisLysAspThrAlaSerThrAsnLeuThrLeuAlaThr 769
QY 2660 ACTGTGATCTTGTTCGTAGTAAACCCGACTGTACGACAACTGCGAATTTAGCGGTGAT 2719
Db 770 AlaProAspIleValArgSerAsnProAspCysThrAlaSerLeuLeuValSerProThr 789
QY 2720 TCT-----TGGAAAACCTTCGGTACGAATTTGCAAGACAAAGCTTTAGCTTCGTGCA 2773
Db 790 SerAlaValTrpValThrLysAlaAsnLeuAlaArgHisAlaPheIleLeuGlnAla 809
QY 2774 GGAACCATTTTTCGTTTAACTCAAAATTTGAAGCCTTTAGCCAAATTTCTTTGAATTG 2833
Db 810 GlyAsnThrLeuAlaLeuThrArgAsnThrGluLeuPheSerGlnPheGlyPheGluLeu 829
QY 2834 CGTGGGTCTATCTCGCAATTTACAATGTAGACTTTAGGACAAATAATCAATTC 2884
Db 830 ArgGlySerCysArgThrThrAsnIleAspLeuGlySerLysIleGlnPhe 846

RESULT 6
Q8VU49 PRELIMINARY; PRT; 602 AA.
AC Q8VU49;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative polymorphic membrane protein (fragment).
OS Chlamydia psittaci (Chlamydia psittaci).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_taxID=83554;
RN [1]
RP SEQUENCE FROM N.A.
RC Laroucau K., Souriau A., Rodolakis A.;
RT "Isolation of a new pmp sequence and evidence of pmp polymorphism in
RT serotype-1 Chlamydia psittaci strains."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF243418; AAL36962.1; -.
DR InterPro; IPR006315; Autotransport.
DR InterPro; IPR003368; Chlamydia_pmp.
DR Pfam; PF02415; DUF145; 1.
DR TIGRFAMs; TIGR01414; autotrans_bar1; 1.
DR NON_TER 1
FT SEQUENCE 602 AA; 65562 MW; CA486CFACEC131E2 CRC64;
SQ SEQUENCE 602 AA; 65562 MW; CA486CFACEC131E2 CRC64;

Alignment Scores:
Pred. No.: 1,08e-64 Length: 602
Score: 1237,00 Matches: 261
Percent Similarity: 59.25% Conservative: 104
Best Local Similarity: 42.37% Mismatches: 205
Query Match: 22.87% Indels: 46
DB: 2 Gaps: 14

US-09-428-122-1 (1-3000) x Q8VU49 (1-602)
QY 1130 GGTGATGCTGCTTTTGTAGGAATACAGTCACTTCTACTCTCGGACGAAT----- 1183
Db 2 GlyAspIleThrPheAspGlyAsnLysIleThrThrSerArgSerSerThrVal 21
QY 1184 ---AGAACTAGTATCGACTTAGGACGAGTGAAGATGACAGCTTTGCTGCTCTGCT 1240
Db 22 LysArgAsnSerIleSerLeuGlySerGlyLysPheThrLysLeuAsnAlaLysGlu 41
QY 1241 GGTAGAGCCATCTACTTCTATGATCCCATATA---ACTACAGGATCATCCACACAGTTACA 1297
```

Db 42 GlyPheGlyIlePhePheTyrAspProIleAlaAsnThrGlyAspThrAsnThrGluIle 61
QY 1298 GATGCTTTAAAGTTAATGAGACTCCGCGAGATTCTGCACATACATAACAGGGAACATC 1357
Db 62 GluLeuAsnLysAlaGlu-----GlyGlySerThrThrTyrThrGlyLysIle 77
QY 1358 ATCTTCACAGAGAAAGTTATCAGACAGAGAGCGCGAGATTCTPAAAAATCTTACTTCG 1417
Db 78 ValPheSerGlyGluLysLeuSerAspGluLysLysValAlaAspAsnLeuLysSer 97
QY 1418 AAGTACTACAGCTGTAACTCTTTCAGGAGTACTCTATCTTTAAACATGCGAGTACT 1477
Db 98 TyrPheThrGlnProLeuLysIleGlyAlaGlySerLeuValLeuLysAspGlyValThr 117
QY 1478 CTGCAGACTCAGGCATCTCACTCAACAGCGAGATTCTCGTCTCGAATAGCAGCTAGAACT 1537
Db 118 LeuGluAlaLysLysValSerGlnThrAspGlySerThrValValMetAspLeuGlyThr 137
QY 1538 ACTTAGAACCTGCT-----GATACAGACATCAAAACAATTTGGTCATTAACATC 1588
Db 138 ThrLeuGlnThrSerSerGlyGluThrIleThrLeuThrAsnLeuAspIleAsnVal 157
QY 1589 ACTTCTATAGCGGTGCA-----AGAAGGCAAAATAGAACCAAGCTAGC 1636
Db 158 AlaserLeuGlyGlyGlyValAlaProAspProAlaLysValGluAlaGlnAlaser 177
QY 1637 TCMAAAATCTGACTTATCTGGAACCATCTTATTCGACCGCGACGGCGACGTTTAT 1696
Db 178 GlyLysThrValThrIleAsn--AlaValAsnLeuValAspThrAspGlyAsnAlaTyr 196
QY 1697 GAAATCATPAGTTTAAAGAACTCTCAGCTCCTAGCATCTTAGAGCTCAAGCTCTGGA 1756
Db 197 GluTyrProIleLeuAlaThrSerGlnProPheThrAlaIleAlaLysAlaGlySer 216
QY 1757 ACTGTAACAGACCGCAGTCACTCCAGAT-----CCTATAATGGGT 1798
Db 217 SerGlyThrThrThrThrProThrAspAsnLeuLysAsnTyrThrProThr----- 234
QY 1799 GAGAAATTCATACCGCTATCAGGGAACCTTGGGCGCCCAATTGTTTGGGGACAGGGCT 1858
Db 235 -----HisTyrGlyTyrGlnGlyAsnTrp--ThrValThrTrpLysLeuGlyThr 250
QY 1859 TCTACG-----ACTGCAACCTTCAACTGAGCTAAACCTGGCTATATCTCAATCCC 1909
Db 251 SerAlaGlnGluGluThrAlaThrLeuThrTrpGluGlnThrAspTyrSerProAsnPro 270
QY 1910 GACGCTATCGGCTCTTATAGTCCCTATAGCTTATGAAATGCAATTTATAGATATGCTCT 1969
Db 271 GluArgGlnGlyProLeuValProAsnThrLeuTrpGlySerPheSerAspIleArgAla 290
QY 1970 CTCATTATCTTATGAGACTGCAACGAGGCTTGCAGGAGACCGTGTCTTTTGGTGT 2029
Db 291 IleGlnAsnLeuIleAspIleSerValAsnGlyAlaAspTyrArgGlyPheTrpVal 310
QY 2030 GTGGATTATCTAATCTTTCCATAGGATAGTACAAAACACAGACGCGGTTCCGCAT 2089
Db 311 SerGlyLeuGlyAsnPheLeuHisLysSerGlySerAsnThrLysArgLysPheArgHis 330
QY 2090 TTGAGTGGCGGTATGTCATAGAGGAAACCTACATCTTGTTCAGATAAGATCTTAGT 2149
Db 331 HisSerAlaGlyTyrAlaLeuGlyValTyrAlaGlnThrSerThrGluAspValPheSer 350
QY 2150 GCTGCATTTGTGAGCTCTTTGGAAGAGATAGACTACTTTGTAGCTAAGAACTCAAGGT 2209
Db 351 AlaAlaPheCysGlnLeuPheGlyLysAspLysAspTyrPheValSerLysAsnSerSer 370
QY 2210 ACAGTCTACGAGGAACCTCTTATACAGACACAAACCTATATCTCTCTCTCTTCTGTC 2269
Db 371 AsnIleTyrAlaGlySerIleTyrTyrGlnHisIleSerTyrTyrAsnAlaTrpGlnAsn 390
QY 2270 AAATCAGCGCTTGTGCTTATAGTCTTCTACAGAGATCTCTGCTCTCTTTCAGGA 2329
Db 391 LeuLeuGln-----SerThrIleGlyAlaGluAlaProLeuValLeuAsnAla 406

QY 2330 AACCTTAGCTACACCCATACGATACGATCTGAAACCAAGTATACA----- 2377
Db 407 GlnLeuThrTyrCysHisAlaSerAsnAsnMetLysThrAsnMetThrAsnThrTyrVal 426
QY 2378 -----ACATATCTCTACTGTTTAAAGAGCTGGGGGAATGATAGTTTCGCTTTA 2425
Db 427 ProLysAsnValThrLeuSerGluIleLysGlyAspTrpGlyAsnAspCysPheGlyVal 446
QY 2426 GAATTCGGTGGAGAGCTCCGATTTCGTAGATGAAAGTGTCTTATTGACAGCTACATG 2485
Db 447 GluPheGlyAlaMetAlaProIleGluThrProSerSerPheLeuPheAspArgTyrSer 466
QY 2486 CCTTCATGAATTCAGTTTGTCTATGCACATCAGGAAGGTTTTAAAGACACGGGAACA 2545
Db 467 ProPheLeuGlnLeuGlnLeuValHisAlaHisGlnAspAspPheLysGluAsnAsnSer 486
QY 2546 --GAAGCTCGTGAATTGGGAAGTAGCGCTCTGTGAATCTTGCTTACCTATCGGGATC 2602
Db 487 AspGlnGlyArgTyrPheGluSerSerAsnLeuThrAsnLeuSerMetProIleGlyIle 506
QY 2603 CGATTGATAAGGAATCAGACTGCCAAGATGCAACGCTACAACTCACTCTTGGTTACT 2662
Db 507 LysPheGluArgPheAlaTyrAsnAspValAlaSerTyrHisLeuThrAlaAlaTyrAla 526
QY 2663 GTGGATCTTCTGTAGTAACCCGACTGTAGCACAACACTGCGAATTAGCGGTGATTCT 2722
Db 527 ProAspIleValArgSerAsnProAspCysThrAlaSerLeuLeuValSerProThrSer 546
QY 2723 -----TCGAAAACCTTCGCTACGAAATTTGGCAGACAAAGCTTTAGTCTCTCGTCAGGG 2776
Db 547 AlaValThrValThrLysAlaAsnAsnLeuAlaArgSerAlaPheMetLeuGlnAlaGly 566
QY 2777 AACCAATTTTGTCTTAACTCAAAATTTGAAGCTTTAGCCAAATTTCTTTTGAATTCGT 2836
Db 567 AsnTyrLeuAlaLeuSerHisAsnMetGluLeuPheSerGlnPheGlyPheGluIleArg 586
QY 2837 GGGTATCTCCGAATTAACATGTAGCTAGGACGAGCAAAATACCAATTC 2884
Db 587 GlySerSerArgThrTyrAsnValAspLeuGlySerLysIleGlnPhe 602

RESULT 7

Q8VL57 PRELIMINARY; PRT; 601 AA.
ID Q8VL57; AC
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative polymorphic membrane protein (Fragment).
OS Chlamydia psittaci (Chlamydia phila psittaci).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phila.
OX NCBI_Taxid=83554;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POS, and LUG;
RA Laroucau K., Souriau A., Rodolakis A.;
RT "Isolation of a new pmp sequence and evidence of pmp polymorphism in serotype-1 Chlamydia psittaci strains."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF243416; AAL36960.1; --
DR EMBL; AF243417; AAL36961.1; --
DR InterPro; IPR006315; Autotransport.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF02415; DUF145; 1.
DR TIGRFAMs; TIGR01414; autotrans_bar1; 1.
FT NON_TER 1 1
SQ SEQUENCE 601 AA; 65476 MW; D6AA97EC9072C757 CRC64;

Alignment Scores:

Pred. No.: 1.15e-63 Length: 601
Score: 1219.50 Matches: 259
Percent Similarity: 59.25% Conservative: 106
Best Local Similarity: 42.05% Mismatches: 204

Query Match: 22.55% Indels: 47
 DB: 2 Gaps: 15

US-09-428-122-1 (1-3000) x Q8VL57 (1-601)

QY 1130 GGTGACATTGCTTTTATAGGAATACAGTCACTTCTACTACTCTCTGGGACGAAT----- 1183
 |||||
 DB 2 GlyAspIleThrPheAspGlyAsnLysIleIleThrThrSerArgSerSerSerThrVal 21
 |||||
 QY 1184 --AGAAGTAGTATCGACTTATAGGAACGAGTGCAGAGATGACAGCTTTGCGTTCGCTGCT 1240
 |||||
 DB 22 LysArgAsnSerIleSerLeuGlySerGlyGlyLysPheThrLysLeuAsnAlaLysGlu 41
 |||||
 QY 1241 GGTAGAGCCATCTACTTCTATGATCCCATN---ACTACAGGATCATCCACAACAGTTACA 1297
 |||||
 DB 42 GlyPheGlyIlePhePheThrAspProIleAlaAsnThrGlyAspThrAsnThrGluIle 61
 |||||
 QY 1298 GATGCTTTAAAGTTAATGAGACTCCGCGAGATTCGCACTACATACATACAGGGAACATC 1357
 |||||
 DB 62 GluLeuAsnLysAlaGlu-----GlyGlySerThrThrTyThrGlyLysIle 77
 |||||
 QY 1358 ATCTTTCAGAGAGAAAAGTTATCAGAGACAGAGCGCGAGATTCCTAAATCTTACTTTCG 1417
 |||||
 DB 78 ValPheSerGlyGluLysLeuSerAspGluGluLysLysValAlaAspAsnLeuLysSer 97
 |||||
 QY 1418 AAGCTACTACAGCTGTAACTCTTTCAGGAGTACTCTATCTTTAAACATCGAGTGAAT 1477
 |||||
 DB 98 TyrPheThrGlnProLysIleGlyAlaGlySerLeuValLeuLysAspGlyValThr 117
 |||||
 QY 1478 CTGCAGACTCAGGCATTCCTCAACAGGAGATTCCTCTCGAATCGAGCTAGGAAT 1537
 |||||
 DB 118 LeuGluAlaLysLysValSerGlnThrAspGlySerThrValValMecAspLeuGlyThr 137
 |||||
 QY 1538 ACTCTAGAACCTGCT-----GATACCTAGCACCATAAACAAATTTGGTTCATTAACTC 1588
 |||||
 DB 138 ThrLeuGlnThrSerSerGlyGluThrIleThrLeuThrAsnLeuAspIleAsnVal 157
 |||||
 QY 1589 AGTCTATAGACGGTGCA-----AAGAAGCGCAAAATAGAACCAAGCTACG 1636
 |||||
 DB 158 AlaSerLeuGlyGlyGlyValAlaProAspProAlaLysValGluAlaGlnAlaSer 177
 |||||
 QY 1637 TCRAAAATCTGACTTTATCTGGAACCATCACTTTATGGACCCGCGGACGCTTTAT 1696
 |||||
 DB 178 GlyLysThrValThrIleAsn---AlaValAsnLeuValAspThrAspGlyAsnAlaThr 196
 |||||
 QY 1697 GAAATCATAGTTAAGAAATCCTCAGTCTCAGCATCTTAGAGCTCAAGCTCTCTGGA 1756
 |||||
 DB 197 GluThrProIleLeuAlaThrSerLysProPheThrAlaIleAlaLysAlaGlySer 216
 |||||
 QY 1757 ACTGTAACAGCACCGCAGTGAATCCAGAT-----CCTATAATGGGT 1798
 |||||
 DB 217 SerGlyThrThrThrThrProThrAspAsnLeuLysAsnThrThrProThr----- 234
 |||||
 QY 1799 GAGAAATTCANTATCGCTATCAGGAATCTGGGCCCAATTTGTTGGGGACAGGGCT 1858
 |||||
 DB 235 -----HisThrGlyThrGlnGlyAsnTrp---ThrValThrTrpLysLeuGlyThr 250
 |||||
 QY 1859 TCTACG-----ACTGCAACCTTCAACTGGACTTAAACCTGGCTATATCTCTAATCCC 1909
 |||||
 DB 251 SerAlaGlnGluGluThrAlaThrLeuThrTrpGluGlnThrAspThrProAsnPro 270
 |||||
 QY 1910 GAGCGTATCGCTCTTTAGTCCATAATAGCTTATGGAATGCATTTATAGATTTAGCTCT 1969
 |||||
 DB 271 GluArgGlnGlyProLeuValProLysThrLeuTrpGlySerPheSerAspIleArgAla 290
 |||||
 QY 1970 CTCCTATCTTATGAGACTGCAACGAGGTTCCAGGAGACCGCTCTTTGCTGT 2029
 |||||
 DB 291 IleGlnAsnLeuIleAspIleSerValAsnGlyAlaAspThrArgArgGlyPheTrpVal 310
 |||||
 QY 2030 GCTGGATTATCTAATCTTCTCCATAGGATAGTACAAAACACGACGCGGGTTCCGCAT 2089
 |||||
 DB 311 SerGlyLeuGlyAsnPheLeuHisLysSerGlySerAsnThrLysArgLysPheArgHis 330
 |||||

QY 2090 TTGAGTGGCGTTTATGTCATAGCAGGAAACCTACATACTTGTTCAGATAAGATCTTGT 2149
 |||||
 DB 331 HisSerAlaGlyThrAlaLeuGlyValThrAlaGlnThrSerThrGluAspValPheSer 350
 |||||
 QY 2150 GCTGCATTTTGTGCTCTTTTGAAGAGATAGAGACTACTTTGTAGTACAAATCAAGGT 2209
 |||||
 DB 351 AlaAlaPheCysGlnLeuPheGlyLysAspLysAspThrPheValSerLysAsnSerSer 370
 |||||
 QY 2210 ACAGTCTACGGAGAACTCTATTACAGACACACAAACCTATATCTCTCTCTCTTC 2269
 |||||
 DB 371 AsnIleThrAlaGlySerIleThrGlnHisIleSerThrThrAsnAlaThrGlnAsn 390
 |||||
 QY 2270 AAATCTACGGCTTGTCTGTCTATGTTCTCTACAGAGATTCCTGTCTCTCTCTACGGA 2329
 |||||
 DB 391 LeuLeuGln-----SerThrIleGlyAlaGluAlaProLeuValLeuAsnAla 406
 |||||
 QY 2330 AACCTTAGCTACACCCATACGGATACAGTCTGAAACCCAAAGTATACA----- 2377
 |||||
 DB 407 GlnLeuThrThrCysHisAlaSerAsnAsnMetLysThrAsnMetThrAsnThrTy-Val 426
 |||||
 QY 2378 -----ACATATCTCTGTTTAAAGAGCTGGGGGATGATAGTTTCGCTTTA 2425
 |||||
 DB 427 ProLysAsnValThrLeuSerGluIleGlyAspTrpGlyAsnAspCysPheGlyVal 446
 |||||
 QY 2426 GAATTCGGTGAAGAGCTCCGATTTGCTTAGATGAAAGTCTCTATTGACAGCTACATG 2485
 |||||
 DB 447 GluPheGlyAlaMetAlaProIle---GluAsnAlaSerPheLeuPheAspArgTyrSer 465
 |||||
 QY 2486 CCCTTCATGAAATTCGATTTGCTCATGCACATCAGGAAGGTTTTTAAAGACAGGGAACA 2545
 |||||
 DB 466 PropheLeuGlnLeuGlnLeuValHisAlaHisGlnAspAspPheLysGluAsnAsnSer 485
 |||||
 QY 2546 ---GAAGCTCGTGAATTTGGAGTAGCGCTCTTGTGAATCTTGCCTTACCTATCGGATC 2602
 |||||
 DB 486 AspGlnGlyArgTyrPheGluSerSerAsnLeuThrAsnLeuSerMetProIleGlyIle 505
 |||||
 QY 2603 CGATTGTATAGGAATCAGACTGCCAAGTCAACATCTCAACTCTTGTGTTATACT 2662
 |||||
 DB 506 LysPheGluArgPheAlaThrAsnAspValAlaSerThrHisLeuThrAlaAlaThrAla 525
 |||||
 QY 2663 GTGGATCTTGTTCGTAGTAAACCCGACTGTACGACACACTCGCAATTTAGCGGTATCT 2722
 |||||
 DB 526 ProAspIleValArgSerAsnProAspCysThrAlaSerLeuLeuValSerProThrSer 545
 |||||
 QY 2723 -----TGAAAACTTCGGTACGAAATTTGCAAGACAGCTTGTAGTCTCTCGTCAGGG 2776
 |||||
 DB 546 AlaValThrValThrLysAlaAsnLeuAlaArgSerAlaPheMetLeuGlnAlaGly 565
 |||||
 QY 2777 AACCATTTTGTCTTAACTCAAAATTTGAAGCTTTAGCAATTTCTTTTGAATTCGT 2836
 |||||
 DB 566 AsnThrLeuAlaLeuSerHisAsnMetGluLeuPheSerGlnPheGlyPheGluIleArg 585
 |||||
 QY 2837 GGCTCATCTCGCAATTAACAATGTAGCTTAGGACGAGCAAAATCCAAATC 2884
 |||||
 DB 586 GlySerSerArgThrThrAsnValAspLeuGlySerLysIleGlnPhe 601
 |||||

RESULT 8
 Q8VU50 PRELIMINARY; PRT; 700 AA.
 ID Q8VU50;
 AC Q8VU50;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DE Putative polymorphic membrane protein (Fragment).
 OS Chlamydia psittaci (Chlamydia phila psittaci).
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phila.
 OX NCBI_TaxID=83554;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=POS;
 RA Laroucau K., Souriau A., Rodolakis A.;
 RT "Isolation of a new pmp sequence and evidence of pmp polymorphism in
 serotype-1 Chlamydia psittaci strains.";

Wed Dec 17 09:17:08 2003

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF243415; AAL36959.1; -
 DR InterPro: IPR003368; Chlamydia_PMP.
 DR Pfam: PF02415; DUF145; 1
 DR TIGRFAMs: TIGR01376; POMP_repeat; 3.
 FT NON_TER 700
 SQ SEQUENCE 700 AA; 74200 MW; 2C38E504F48D1D1 CRC64;

Alignment Scores: 4,48e-61 Length: 700
 Pred. No.: 1175.50 Matches: 291
 Score: 1175.50
 Percent Similarity: 50.00% Conservative: 123
 Best Local Similarity: 35.14% Mismatches: 237
 Query Match: 21.73% Indels: 177
 DB: 2 Gaps: 2

US-09-428-122-1 (1-3000) x Q8VU50 (1-700)

QY 89 ATATATTTCACATGAAGTCTCTTCCCAAGTTTGTATTCTACATTTGCTATTTC 148
 Db 5 ValTyrTrpPheLeuSerSer-----SerLeuLeuAlaSerAsn 18
 QY 149 CCTTGTCTATGATT---GCTACCGAGACAGTTTGGATTCAAGTGGAGTTTCGATGGG 205
 Db 19 SerLeuSerPheAlaGlnValThrAsnGluThrLeuThrSerSerSerTyrAsnGly 38
 QY 206 AAT---AAAAATGGTAAATTTTCAGTTCGTGAGAGTCAGGAAGATGCTGAACCTAC 262
 Db 39 AsnValThrSerAspGluPheGluValLysGluThr-----ThrSerGlyAlaLeuTyr 56
 QY 263 CTATTTAAGGAATATGCTACTCTAGAAAATATCTCTGGAACAGCAGCAATACAAAA 322
 Db 57 ThrCysGluGlyAsnValCysIle---SerTyrAlaGlyLysAspSerProLeuAsnLys 75
 QY 323 AGCTGTTTAAACACTAAGGGGATTTGACTTTTCAGGTAAAGGAACTCTCTATTG 382
 Db 76 SerCysPheSerGluThrThrGluAsnLeuSerPheIleGlyAsnGlyTyrThrLeuCys 95
 QY 383 TTCCAAACGGTGGATGTCAGGAGCTGTAGCAGGGGCTGCTGTTAACAGCAGGCTGGTAG 442
 Db 96 PheAspAsnIleThrThrAlaSerAsnProGlyAlaIleAsnValSerGlyAspGln 115
 QY 443 AAATCTACAGCTTATAGGTTTCTCGCTATCTTTATTCGCTCTCTCGGAAGTCG 502
 Db 116 LysThrLeuAsnValSerGlyPheSerLeuPheSerCysAlaHisCysProGly--- 134
 QY 503 ATACTACCGCAAGGAGCGGTAGCTGCTCTACGGGTAGCTTGAGTTTGACAAAAAAT 562
 Db 135 ---ThrThrGlyTyrGlyAlaIleGln-----ThrLysGly 145
 QY 563 GTCAGTTTGTCTCTTCAGCAAAACTTTTCAACGGGATAATGCGGTGCTATCCCGCAAA 622
 Db 146 ValSer----- 147
 QY 623 ACTCTTTCATTAACAGGACTACATGTCTGTTTCTGAAATACCTCTCCTCAAG 682
 Db 147 ----- 147
 QY 683 AAAGGGGAGCCATTTCAGACTTCGATGCCCTTACCATTACTGTGAACCAAGGGGAAGTC 742
 Db 147 ----- 147
 QY 743 TCTTTTCTGCAATATCTTCTCGGATTCAGAGCTGCAATTTTACAGAGCCTCGGTG 802
 Db 147 ----- 147
 QY 803 ACTATTCTAATAGCTAAAGTTTCCTTTATTTGACAAATAAGTTCACAGGCGAGCTCC 862
 Db 148 ThrPheSerGlyAsnAsnLysLeuIlePheAspAsnAsnCysSerThrGly----- 164
 QY 863 TCACAAACGGGATATGTCAGGAGTGTATC---TGTGCTTATAAACTAGTACAGAT 919
 Db 165 -----GluGlyGlyAlaIleLysCysAla-----ThrGlySerAsn 176

QY 920 ACTAAGTCAACCTCACTGGAATCAGATGTTACTCTTACAGCAACAATACATACGACAAC 979
 Db 177 AlaGluLeuLysLeuGluGlyAsnSerTyrValValPheSerGlyAsnSerSerGlnLys 196
 QY 980 GCGGAGGAGCTATCTATGTGAAAAAGCTCGAAGTGGCTTCGGAGGACTTACCCCTATTC 1039
 Db 197 LysGlyGlyAlaIleTyrThrLysLysLeuThrIleThrAlaAspGlyProThrLeuPhe 216
 QY 1040 AGTAGAATAGTGTCAATGGAGGTACAGCTCTTAAGGTGGAGCCATAGCTACGAAGAT 1099
 Db 217 SerAsnAsnSerValSer---AlaSerSerProLysGlyGlyAlaIleCysLeuAspAsp 235
 QY 1100 ---AGTGGGAGTTCGATTTATCCCGCATAGTGTGTGACATTCCTTTTACGGGAATACA 1156
 Db 236 ThrSerSerGluCysSerLeuThrAlaAsnLeuGlyAspIleThrPheAspGlyAsnLys 255
 QY 1157 GTCACCTCTACTACTCTCTGGGACGAAT-----AGAAGTAGTATCGACTTAGAAGC 1207
 Db 256 ValIleLysThrAsnGlyGlySerSerThrValLysArgAsnAlaIleAspLeuGlySer 275
 QY 1208 AGTGCMAAGATGACAGCTTTCGCTGCTGCTGTAGAGCCATCTACTCTTATGATCCC 1267
 Db 276 GlyGlyLysPheThrLysLeuAsnAlaLysGluGlyPheGlyIlePhePheHisAspPro 295
 QY 1268 ATA---ACTACAGGATCATCCACACAGTTACAGATGCTTAAAGTAAATGAGACTCCG 1324
 Db 296 IleAlaAsnThrGlyGlySerThrGluIle-----GluLeuAsnLysThrGlu 311
 QY 1325 GCAGATTCTGCATCAATATACAGGAACATCATCTTACAGGAGAAAAGTTATCAGAG 1384
 Db 312 SerAspThrThr-----TyrThrGlyLysIleValPheSerGlyGlyLysLeuSerAsp 329
 QY 1385 ACAGAGCCCGCAGATTCTAAATCTTACTTGAAGCTACTACAGCTGTAACCTCTTCA 1444
 Db 330 GluGluLysThrValProAspAsnLeuLysSerTyrPheLysGlnProLeuLysIleGly 349
 QY 1445 GGAAGTACTCTATCTTTAAACATGGAGTACTCTGCAGACTCAGCATTCACTCAACAG 1504
 Db 350 AlaGlySerLeuValLeuLysAspGlyValThrLeuGluAlaLysLysIleThrGlnThr 369
 QY 1505 GCAGATTCTCGTCTCGAAATGGAGTACTAGTAACCTCTAGAA-----CCTGCT 1552
 Db 370 LysGlySerThrValValMetAspLeuGlyThrThrLeuGlnThrProSerSerSerGly 389
 QY 1553 GATCTAGCACCATAACAATTTGCTTAAACATCAGTCTCTATA-----GACGGT 1603
 Db 390 GluThrIleThrLeuThrAsnLeuAspIleAsnIleAlaSerLeuGlyGlyGlyGly 409
 QY 1604 GCAAGAAGGCAAAATAGAAACCAAGCTACGTCAAAAAATCTGACTTTTATCTGGAACC 1663
 Db 410 ThrAlaProAlaLysLeuAlaThrAsnThrAlaSerGlnAlaIleSerIleAla---Ala 428
 QY 1664 ATCACTTTATGGACCGGACCGGACGCTTTTATGAAAATCATAGTTTAAAGAAATCCTCAG 1723
 Db 429 ValAsnLeuValAsnThrAspSerAsnThrTyrGluAspProIleLeuSerAlaSerLys 448
 QY 1724 TCCATCAGCATCTTAGAGCTCAAGCTTCTGGAAGTCTTAACAACGCGAGTCACTCCA 1783
 Db 449 SerPheSerAlaIle-----ThrAlaThrThrSerSerSerThrValThrPro 464
 QY 1784 GATCCTATATGGGTGAGAAATTC-----CATTCAGGCTATCAGGGAACTGG 1831
 Db 465 ProGluThrAsnLeuLysAsnTyrThrProProThrHisTyrGlyTyrGlnGlyAsnTrp 484
 QY 1832 GGCCTCAATTTGTTGGGACAGGGGCTTCT-----ACGACTGCAACCTTCACTGG 1882
 Db 485 ---ThrLeuAlaTrpProProGlyGluThrMetGlnLeuLysThrAlaThrLeuAsnTrp 503
 QY 1883 ACTAAACTGGCTATATCTTAATCCCGAGGTATCGGCTCTTTAGTCCCTTAATAGCTTA 1942
 Db 504 GluGlnThrGlyTyrSerProAsnProGluArgValGlySerLeuValProAsnThrLeu 523

```

QY 1943 TGGAAATCATTATAGATATAGTCTCTCCATTATCTTATGAGACTGCAACGAGGG 2002
DB 1943 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
524 TrpGlySerPheSerAspIleArgAlaIleGlnAsnLeuMetAspValSerValAsnGly 543
QY 2003 TTGCAGGAGACCGTCTTTTGTGTGCTGATTAATCTTAACTTCTTCCATAAGGATAGT 2062
DB 2003 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
544 AlaAspTyrSerArgGlyPheTrpValSerGlyLeuAlaAsnPheLeuAsnLysSerGly 563
QY 2063 ACAAAAACACAGACCGGGTTTCGCCATTGTAGTCGGCGGTATATGTCATAGGAGAAACCTA 2122
DB 2063 :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
564 SerAspThrLysArgLysPheArgHisAsnSerAlaGlyTyrAlaLeuGlyValTyrAla 583
QY 2123 CATACTTGTTCAGATAAGATCTTAGTCTCCATTGTGTAGCTCTTGTGGAGAGATAGA 2182
DB 2123 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
584 GlnThrProSerGluAspIlePheSerAlaAlaPheCysGlnLeuPheGlyLysAspLys 603
QY 2183 GACTACTTGTAGTACGAATCAAGGTACAGTCTACGGAGAACTCTCTATTACCGACAC 2242
DB 2183 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
604 AspTyrPheLeuSerLysAsnSerSerThrIleTyrAlaGlySerIleTyrTyrGlnHis 623
QY 2243 AACGAAACCTATATCTCTCTCTTCTTGCAACTACGGCGCTTTCGTTGCTCTTAT 2296
DB 2243 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
624 -----IleSerTyrTrpAsn 628
QY 2297 -----GTTCTACAGAGATTCTGTTCTCTTTTCAGGA 2329
DB 2297 :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
629 AlaTrpGlnAsnLeuLeuGlnAsnThrIleGlyAlaGluAlaProLeuValLeuAsnAla 648
QY 2330 AACCTTAGTACACCCATCAGGATAACGATCTGAAACCAAGTATACA 2377
DB 2330 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
649 GlnLeuThrTyrCysHisAlaSerAsnAsnMetLysThrAsnMetThrAsnThrTyrThr 668
QY 2378 -----ACATATCTCTACTGTTAAAGAAAGCTGGGGAATGATAGTTTCGCTTTA 2425
DB 2378 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
669 ProLysAsnValThrProSerGluIleLysGlyAspTrpGlySerAspCysPheGlyVal 688
QY 2426 GAATTCGGTGAAGACTCCGATT 2449
DB 2426 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
689 GluPheGlyAlaLysAlaProIle 696

```

RESULT 9

```

Q8VU48 PRELIMINARY; PRT; 581 AA.
AC Q8VU48;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Putative polymorphic membrane protein (Fragment).
OS Chlamydia psittaci (Chlamydia phila psittaci).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phila.
OX NCBI_TaxID=833554;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POS;
RA Laroucau K., Souriau A., Rodolakis A.;
RT "Isolation of a new pmp sequence and evidence of pmp polymorphism in
RL serotype-1 Chlamydia psittaci strains.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF243419; AAL36963.1; -.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF02415; DUF145; 1.
FT NON_TER
SQ SEQUENCE 581 AA; 62860 MW; CDDF3C98522E112F CRC64;

```

Alignment Scores:

Pred. No.:	Length:	Matches:
Score:	6,58e-61	581
Percent Similarity:	1172.50	253
Best Local Similarity:	56.30%	Conservative: 100
Query Match:	40.35%	Mismatches: 185
DB:	21.68%	Indels: 89
	2	Gaps: 17

US-09-428-122-1 (1-3000) x Q8VU48 (1-581)

```

QY 1130 GGTGACATGCTCTTTTATAGGAATACAGTCTACTTCTACTACTCTCTGGACGAAT----- 1183
DB 1130 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2 GlyAspIleThrPheAspGlyAsnLysIleIleThrThrSerArgSerSerThrVal 21
QY 1184 ---AGAAAGTAGTATCGACTTTAGGAACAGAGTGCAGGATGACAGCTTTCCTCTGCTGCT 1240
DB 1184 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
22 LysArgAsnSerIleSerLeuGlySerGlyGlyPheThrLysLeuAsnAlaLysGlu 41
QY 1241 GGTAGAGCCATCTACTTCTATGATCCCATATA---ACTACAGGATCATCCCAACAGTATACA 1297
DB 1241 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
42 GlyPheGlyIlePhePheTyrAspProIleAlaAsnThrGlyAspThrAsnThrGluIle 61
QY 1298 GATGCTTAAAGTAAATGAGACTCCGCGAGATCTTCGACTACATATATACAGGAAACATC 1357
DB 1298 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
62 GluLeuAsnLysAlaGlu-----GlyGlySerThrThrTyrThrGlyLysIle 77
QY 1358 ATCTTCACAGAGAAAAGTTTATCAGAGACAGAGCCGAGATTCTTAAATCTTACTTCG 1417
DB 1358 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
78 ValPheSerGlyGlyLysLeuSerAspGluGlyLysValAlaAspAsnLeuLysSer 97
QY 1418 AAGCTACTACAGCTGTAACTCTTTCAGAGGTAATCTATCTTTTAAACATGAGTGAAT 1477
DB 1418 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
98 TyrPheThrGlnProLeuLysIleGlyAlaGlySerLeuValLeuLysAspGlyValThr 117
QY 1478 CTGACAGACTCAGGCAATCTCAACAGGAGATCTCTCTCGAATGAGAGTGAAT 1537
DB 1478 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
118 LeuGluAlaLysLysValSerGlnThrAspGlySerThrAlaValMetAspLeuGlyThr 137
QY 1538 ACTCTAGAACCTGCT-----GATACTAGACCAATAAACAATTTGGTCATTAAACATC 1588
DB 1538 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
138 ThrLeuGlnThrSerSerSerGlyGluThrIleThrLeuThrAsnLeuAspIleAsnVal 157
QY 1589 AGTTCTTACACGCTGCA-----AAGAGGCAAAAATAAGAAACCAAGCTACG 1636
DB 1589 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
158 AlaSerLeuGlyGlyGlyValAlaProAspProAlaLysValGluAlaGlnAlaSer 177
QY 1637 TCRAAAAATCTGACTTTATCTGGAAACCATCTATTATGACCGGACGCGGACGCTTTAT 1696
DB 1637 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
178 GlyLysThrValThrIleAsn---AlaValAsnLeuValAspThrAspGlyAsnAlaTyr 196
QY 1697 GAAATCATAGTTTAAAGAAATCTCAGTCCTACGACATCTTAGAGCTCAAGCTCTCGGA 1756
DB 1697 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
197 GluTyrProIleLeuAlaThrSerGlnProPheThrAlaIleIleAlaLysAlaGlySer 216
QY 1757 ACTGTAAACAGACCGCAGTACTCCAGAT-----CCTATAATGGGT 1798
DB 1757 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
217 SerGlyThrThrThrThrProThrAspAsnLeuLysAsnTyrThrProThr----- 234
QY 1799 GAGAAATTCATTTAGCGCTATCAGGAACTTGGGGCCCAATGTTTGGGGACAGGGGCT 1858
DB 1799 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
235 -----HisTyrGlyTyrGlnGlyAsnTrp---ThrValThrTrpLysLeuGlyThr 250
QY 1859 TCTACG-----ACTGCAACCTTCACTGAGTAAACTGCTATATCTTAATCCC 1909
DB 1859 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
251 SerAlaGlnGluThrAlaThrLeuAsnTrpGluGlnThrGlyTyrSerProAsnPro 270
QY 1910 GAGCGTATCGGCTCTTTAGTCCCTAATAGCTTATGGAATGCATTTATAGATATTAGCTCT 1969
DB 1910 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
271 GluArgValGlySerLeuValProAsnThrLeuTrpGlySerPheSerAspIleArgAla 290
QY 1970 CTCCATTATCTTAGGAGACTGCAACAGAGGTTGCGAGGACCGCTCTTTTGGTGT 2029
DB 1970 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
291 IleGlnAsnLeuMetAspValSerValAsnGlyAlaAspTyrSerArgGlyPheTrpVal 310
QY 2030 GCTGATTTATCTACTTCTCCATAAGATAGTACAAAACACGACCGGGTTCGCCAT 2089
DB 2030 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
311 SerGlyLeuAlaAsnPheLeuAsnLysSerGlySerAspThrTyrArgLysPheArgHis 330
QY 2090 TTGAGTGGCGGTTATGTCATAGAGGAAACCTACATACCTTCTTCAGATAAGATTTAGT 2149
DB 2090 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
331 AsnSerAlaGlyTyrAlaLeuGlyValTyrAlaGlnThrProSerGluAspIlePheSer 350

```


QY 2150 GCTGCAATTTGTCAGCTCTTTGGAGAGATAGAGACTACTTTGTAGCTAAGCAATCAAGT 2209
Db 351 AlaAlaPheCysGlnLeuPheGlyLysAspTyrPheLeuSerLysAsnSerSer 370
QY 2210 ACAGTCTACGGAGGAACTCTCTATTACACAGCAACAGAAACCTATATCTCTCTCTGTC 2269
Db 371 ThrIleTyrAlaGlySerIleTyrGlnHis 381
QY 2270 AAACCTACGGCCTTTGTTGCTGCTTAT 2296
Db 382 1leSerTyrTrpAsnAlaTrpGlnAsnLeuLeuGlnAsnThr 395
QY 2297 GTTCTACAGATTCCTGTTCTCTTTCCAGAACTTAGCTACACCCATACGGATAC 2356
Db 396 IleGlyAlaGluAlaProLeuValLeuAsnAlaGlnLeuThrTyrCysHisAlaSerAsn 415
QY 2357 GATCTGAAACCAAGATATACA 2392
Db 416 AsnMetCysThrAsnMetThrAsnThrTyrThrProLysAsnValThrProSerGluIle 435
QY 2393 AAAGGAAGCTGGGGGAATGATTTTCGCTTTAGAAATCGGTGGAGAGCTCCGATTGC 2452
Db 436 LysGlyAspTrpGlySerAspCysPheGlyValGluPheGlyAlaLysAlaProIle 454
QY 2453 TTAGATCAAGTCTCTATTGAGCAGTACATGCCCTTCATGAAATTCGAGTTTGTCTAT 2512
Db 455 GluThrAlaSerLeuLeuPheAspMetTyrSerProPheValLysLeuGlnLeuValHis 474
QY 2513 GCACATCAGGAAGTTTTAAAGAACAGGGAACA 2569
Db 475 AlaHisGlnAspPheLysGluAsnSerAspGlnGlyArgTyrPheGluSerAsn 494
QY 2570 CGTCTTGTAATCTGCTTACTATCGGATCCGATTCGATGATCGTGTGATACCTGCTGATACCCCGAC 2629
Db 495 AsnLeuThrAsnLeuSerMetProIleGlyValLysLeuGluLysPheSerHisGluAsp 514
QY 2630 GATGCAACGTACATCTACTCTGTTGTTACTGTGATCTGCTGATACCTGCTGATACCCCGAC 2689
Db 515 ThrAlaSerTyrAsnLeuThrLeuAlaTyrAlaProAspIleValArgSerAsnProAsp 534
QY 2690 TGTACGACAACTCGCAATTAGCGGTGATTCT 2743
Db 535 CysThrAlaSerLeuLeuValSerProThrSerAlaValTrpValThrLysAlaAsnAsn 554
QY 2744 TTGGGAAGCAAGCTTTAGCTTCGTCGAGGAAACCATTTTGTCTTAACTCAAAATTT 2803
Db 555 LeuAlaArgHisAlaPheIleLeuGlnAlaGlyAsnTyrLeuAla 569
QY 2804 GAAGCCTTTAGCAATTTCTTTTGAATTGCGTGGGTCACTCGCAATTACAATGTAGAC 2863
Db 570 570 -----ThrTyrAsnIleAsp 574

RESULT 10

Q9RB67 ID Q9RB67 PRELIMINARY; PRT; 445 AA.
AC Q9RB67;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Pmp 5.
GN Pmp 5.2.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydiales.
ON NCBI_TaxID=83558;
RX [1]
SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,

RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RL Nucleic Acids Res. 28:2311-2314 (2000).
DR EMBL; AP002545; BAA98231.1; -
DR InterPro; IPR006315; Autotransport.
DR Pfam; PF03797; Autotransporter.
DR TIGRFAMs; TIGR01414; autotrans_bar1; 1.
SQ SEQUENCE 445 AA; 49353 MW; EFA24AFC9C5097A6 CRC64;

Alignment Scores:

Pred. No.: 7,96e-52 Length: 445
Score: 1017.50 Matches: 202
Percent Similarity: 60.99% Conservative: 70
Best Local Similarity: 45.29% Mismatches: 165
Query Match: 18.81% Indels: 9
DB: 16 Gaps: 7

US-09-428-122-1 (1-3000) x Q9RB67 (1-445)

QY 1565 ATAAACAATTGGTTCATTAAACATCAGTTCTATAGACGGTGCAGAGCAAGCAAAATAGAA 1624
Db 3 1leThrAsnLeuSerIleAsnAlaAspThrIleTyrGlyAsnProIleAsnIleVal 22
QY 1625 ACCAAGCTACGTCACAAAATCTGACTTTATCTCGAACCATCACCTTTATTGGACCCGACG 1684
Db 23 AlaSerAlaAlaAsnLysAsnIleThrLeuThrGlyThrLeuAlaLeuValAsnAlaAsp 42
QY 1685 GGCAGCTTTTATGAAATCATAGTTTAAAGAAATCTCAGTCTCAGACATCTTAGAGTCT 1744
Db 43 GlyAlaPheTyrGluAsnHisThrLeuGlnAspSerGlnAspTyrSerPheValLysLeu 62
QY 1745 AAA-----GCTTCTGGAACCTGTAACAGCACCGCAGTGCATCCAGATCCT-----ATA 1792
Db 63 SerProGlyAlaGlyGlyThrIleThrGlnAspAlaSerGlnLysProLeuGluVal 82
QY 1793 ATGGGTGAGAAATTCATTACCGCTATCAGGAACTTGGGGC---CCAATTGTTGGGG 1849
Db 83 AlaProSerArgProHisTyrGlyTyrGlnGlyHisTrpAsnValGlnValIleProGly 102
QY 1850 ACAGGGGCTTCT---ACGACTGCAACCTTCAACTGACTTAAACCTGGCTATATCTCTAAT 1906
Db 103 ThrGlyThrGlnProSerGlnAlaAsnLeuGluTrpValArgThrGlyTyrLeuProAsn 122
QY 1907 CCGAGCGTATCGCTCTTTAGTCCCTATAGCTTATGCAATGCATTTATAGATATAGC 1966
Db 123 ProGluArgGlnGlySerLeuValProAsnSerLeuTrpGlySerPheValAspGlnArg 142
QY 1967 TCTCTCATTTATCTTGGAGACTGCACAAAGAGGTTCGAGGAGACCGTCTTTTGG 2026
Db 143 AlaIleGlnGluIleMetValAsnSerSerGlnIleLeuCysGlnGluArgGlyValTrp 162
QY 2027 TGTGCTGGATTATCTAACTTCTTCCATAAGGATAGTACAAAACACGACGCGGTTTCGC 2086
Db 163 GlyAlaGlyIleAlaAsnPheLeuHisArgAspLysIle---AsnGluHisGlyTyrArg 181
QY 2087 CATTTGAGTGGCGGTTATGTCATAGGAGGAACTTACATCTTGTTCAGATAAGATCTT 2146
Db 182 HisSerGlyValGlyTyrLeuValGlyValGlyThrHisAlaPheSerAspAlaThrIle 201
QY 2147 AGTGTGCTGCTTTTGTGCTGCTTTTGGAGAGATAGACTACTTTGTAGCTAAGATCAA 2206
Db 202 AsnAlaAlaPheCysGlnLeuPheSerArgAspLysAspTyrValValSerLysAsnHis 221
QY 2207 GGTACAGTCTACGAGGAACTCTCTATTACCACACAAACCACTATATCTCTCTCTCT 2266
Db 222 GlyThrSerTyrSerGlyValValPheLeuGluAspThrLeuGluPheArgSer---Pro 240
QY 2267 TGAACACTACGGCTTGTGCTGCTTATGTTCTTCTACAGAGATTCCTGTTCTCTTTCA 2326
Db 241 GlnGlyPheTyrThrAspSerSerSerGlnAlaCysCysAsnGlnValValThrIleAsp 260

```

QY 2327 GGAACCTTAGCTACACCCATACGGATAACAGTCTGAAACCAAGTATGATACACATATCCT 2386
Db 261 MetGlnLeuSerTyrSerHisArgAsnAspMetLysThrTyrThrThrTyrPro 280
QY 2387 ACTGTTAAAGAACTGGGGAATGATAGTTTCCTGCTTTAGAAATTCGGTGGAGAGCTCCG 2446
Db 281 GluAlaGlnGlySerTTPAlaAsnAspValPheGlyLeuGluPheGlyAlaThrThrTyr 300
QY 2447 ATTTGCTTAGAAGAGTCTCTATTGAGCAGTACATGCCCTTCATGAATTCAGTTT 2506
Db 301 TyrTyrProAsnSerThrPheLeuPheAspTyrTyrSerProPheLeuArgLeuGlnCys 320
QY 2507 GTCTATGCATCAGGAGGTTTAAAGAACAGGAGACAGAGCTCGTGAATTTGGAGT 2566
Db 321 ThrTyrAlaHisGlnGluAspPheLysGluThrGlyGlyGluValAlaHisPheThrSer 340
QY 2567 AGCGCTCTTGTAATCTTTCCTTACCTATCGGATCCGATTGATAAGGAATCAGACTGC 2626
Db 341 GlyAspLeuPheAsnLeuAlaValProIleGlyValLysPheGluArgPheSerAspCys 360
QY 2627 CAAGATCAAGCTACATCTACTCTGGTTATCTACTGTGGATCTTGTCTAGTAACCC 2686
Db 361 LysArgGlySerTyrGluLeuThrPheAlaTyrValProAspValIleArgLysAspPro 380
QY 2687 GACTGTACGACACACTCGGAATTAGCGGTGATTCTTGGAAACCTTCGGTACGAATTG 2746
Db 381 LysSerThrAlaThrLeu---AlaSerGlyAlaThrTyrSerThrHisGlyAsnAsnLeu 399
QY 2747 GCAACAGAGCTTTAGTCTCTCGTCGAGGGAACCACTTTTCTGTTTAACCTCAAAATTTGAA 2806
Db 400 SerArgGlnGlyLeuGlnLeuArgLeuGlyAsnHisCysLeuIleAsnProGlyIleGlu 419
QY 2807 GCCTTTAGCCAAATTTCTTTGAATTCGTGGGTGATCTCGCAATTAACAATGTAGACTTA 2866
Db 420 ValPheSerHisGlyAlaIleGluLeuArgGlySerSerArgAsnTyrAsnIleAsnLeu 439
QY 2867 GGAGCAAAATACCAATC 2884
Db 440 GlyGlyLysTyrArgPhe 445

```

RESULT 11

```

P71134
ID P71134 PRELIMINARY; PRT; 649 AA.
AC
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative outer membrane protein (Fragment).
OS Chlamydomophila abortus.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OX NCBI_TaxID=83555;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ovine abortion S26/3;
RA Longbottom D., Russell M., Dunbar S.M., Jones G.E., Herring A.J.;
RT "98kDa protein genes from ovine abortion strain S26/3 Chlamydia
RT psittaci.".
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U72499; AAB18187.1; -.
DR InterPro; IPR006315; Autotransport.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF02415; DUF145; 1.
DR TIGRFAMs; TIGR01414; autotrans_bar1; 1.
FT NON TER 1
SQ SEQUENCE 649 AA; 70091 MW; 13747C68066A7F50 CRC64;

```

Alignment Scores:

```

Score: No.: 5.39e-47 Length: 649
Pct Ident: 935.50 Matches: 217
Percent Similarity: 46.51% Conservative: 103
Best Local Similarity: 31.54% Mismatches: 243
Query Match: 17.30% Indels: 125
DB: 2 Gaps: 12

```

```

US-09-428-122-1 (1-3000) x P71134 (1-649)
QY 1064 ACAGCTCTCTAAAGTGGAGCCATAGCTATCGAAGATAGTGGGAATTGAGTTTATCCGCC 1123
Db 6 ThrAlaProAlaAlaAlaValThrAlaLysGlnProGluAlaSerLeuAlaLys 25
QY 1124 GATAGTGGTGCATTTGCTTTTGGGAATACAGTCTACTCTACTCTCTGGGACGAAT 1183
Db 26 AlaThrSerGlyIleProAlaSerGlySerAlaValSerValProAlaProAlaProThr 45
QY 1184 AGAAGTAGTAGTACCTTAGGAACGAGTGCAGAGTACAGCTTTGCTGCTGGT 1243
Db 46 LysThrProLeuLysIle----- 51
QY 1244 AGAGCCATCTACTCTCTATGATCCATACTACAGGATCATCCACACAGTTTACAGATGTC 1303
Db 52 -----AsnAlaProAspThrGlnAspProGluIle 61
QY 1304 TTAAGAGTTAATGAGCTCCGCGCATTTCTGCATCAATATACAGGGAACATCATCTTC 1363
Db 62 GlnLysValAlaAlaGluAlaGlnGlnSerAlaValTyrAsnGlyLysIleValPhe 81
QY 1364 ACAGAGAAAAGTTATCAGAGACAGCGCGCAGATTCTAAAAATCTTACTTCAAGCTA 1423
Db 82 SerGlyGluLysLeuSerSerGluAspAlaLysAsnProLeuAsnAlaThrSerValIle 101
QY 1424 CTACAGCTGTAATCTTTTCTCAGGAGTACTCTTCTTAAACATGAGTACTCTGCAG 1483
Db 102 HisAsnAspValSerLeuGluAlaGlyThrLeuValLeuSerSerGlyAlaGlyLeu 121
QY 1484 ACTCAGGATTCACCTCAACAGGCGAGATTCTCGTCTCGAAATGAGCTAGGAACTACTCTA 1543
Db 122 ValAspSerPheThrGlnLysGluGlySerLeuIleValMetAspGlyGlyThrSerIle 141
QY 1543 ----- 1543
Db 142 IleThrAsnValThrProAlaSerGluGlyLeuGlnSerArgSerThrProProSerPro 161
QY 1543 ----- 1543
Db 162 LysAsnAlaIleProValIleArgAlaValSerLysValIleAlaSerSerLeuIleAsn 181
QY 1544 -----GAACCTGCTGATCTAGC----- 1561
Db 182 LeuArgGluArgAlaAspSerGlyAlaGlyAlaValValProThrIleGluGluSerPro 201
QY 1562 -----ACCATAAACATTTGGTCAATTAACTACATCAGTTCTATAGACGGTCAAG 1609
Db 202 AspGlySerIleThrIleThrAsnLeuAlaValAsnLeuAspSerLeuGluAsnGlyLys 221
QY 1610 AAGGCAAAATAGAAACCAAGCTACGTCAAAATAATCTGACTTTATCTGGAACCATCACT 1669
Db 222 ValIleThrLeuAlaAlaLysGlySerGlySerValThrLeuThrGlyAspGln 241
QY 1670 TTAITGGACCGCGGCGACGCTTTATGAAATCATAGTTTAAAGAAATCTCAGTCCCTAC 1729
Db 242 PheGlnAspSerSerGlnAsnPheTyrAspAsnProLeuLeuAsnLysAsnPheThrLeu 261
QY 1730 GACATCTTAGAGCTCAAGACTTCTGGAACCTGTAACAGCACCGCAGTACTCCAGATCCT 1789
Db 262 AsnPheLeuAspIleSerAla-----ProAsp--- 270
QY 1790 ATAATGGGTGAGAAATTCAT----- 1810
Db 271 -----AlaGluLysIleHisThrGluGlyPheAsnIleIleProGlnGlyAlaThrSer 288
QY 1811 -----TACGCTATCAGGAACTTGGGGCCCAATT---GTTGGGGACAGGGCTTCT 1861
Db 289 SerAsnLeuGlyTyrGlnGlyLysTyrGluValThrGluValLysAspSerSerGlyLys 308
QY 1862 ACGACTCAACCTTCACTGAGCTAAACTGGCTATATTCCTTAATCCCGAGCGTATCGGC 1921
Db 12

```

```
Db 309 ValSerPheGluMetLysTrpValSerThrGlyTyrIleProThrAlaAsnArgAla 328
QY 1922 TCCTTAGTCCCTAATAGCTTATGGAATGCAATTTATAGATATTAGCTCTCTCCATTTATCTT 1981
DB 329 ThrLeuValProAsnSerValTrpCysSerAlaIleAspMetArgAlaPheGlnAsnLeu 348
QY 1982 ATGGAGACTGCAAAAGAGGGTTGAGGAGACCGTCTTTTGTGTGGTGGATATCT 2041
DB 349 ValGluValSerThrGluGlyAspPheHisArgGlyLeuTrpIleSerGlyIleSer 368
QY 2042 RACTCTTCCATAAGATAGTACAAAACACAGCAGCGGTTTCGCCATTTAGTGGCGGT 2101
DB 369 AsnPhePheHisLysAspSerThrLysValGlnGluGlyPheArgHisIleSerSerGly 388
QY 2102 TATGTCATAGGAGAAACCTACATCTTGTTCAGATAAGATTCTTAGTGTCTCATTTGT 2161
DB 389 TyrValValGlyValSerThrGlnProIleSerAsnLysValMetAspLeuAlaPheCys 408
QY 2162 CAGCTCTTTGGAAGATAGAGACTACTTTGTAGCTAAGATCAAGGTACAGTCTACGA 2221
DB 409 GlnMetLeuGlySerLysAspTyrArgLeuAlaAspAlaArgSerHisValTyrAla 428
QY 2222 GGAACCTCTATTACAG-----CACACGAAACCTATATCTCTCTCTCT 2266
DB 429 AlaSerIleHisThrLysCysGluLysLeuValAsnHisTyrThrPhe----- 444
QY 2267 TGCAAACTACGGCCTTGTGTTGCTTATGTTCTTACAGATTCCTTTCTCTTTCA 2326
DB 445 SerLysArgLysGlyAlaIleLeuAlaArgLysProGluLysSerProIleIlePheAsp 464
QY 2327 GGAACCTTAGCTACACCATACGGATAACGATCTGAAACCAAGTATACACATATCT 2386
DB 465 AlaGlnLeuSerTyrSerLeuSerHisAsnSerMetThrThrLysHisThrProAsnPro 484
QY 2387 ACTGTTAAAGAGAGCTGGGGAGATAGTTCCTGCTTTAGAAATTCGGTGGAAAGCTCG 2446
DB 485 SerSerArgGlyLysTrpAsnAsnHisCysValAlaGlyGluLeuGlySerTyrLeuPro 504
QY 2447 ATTTGCTTAGATCAAAAGTCTCTATTGTAGCAGTACATCCCTTCANGAAATTCAGTTT 2506
DB 505 IleLeuValAspHisProAlaIle---GluGluLeuPheProPheValLysLeuHisIle 523
QY 2507 GTCTATGTCATCAGAGAGGTTTAAAGAACAG-----GGAACAGAGCTCGTGAATTT 2560
DB 524 ValPheValGlnGlnGluAspPheLysGluThrGlnGlyGlyThrGluAsnArgAsnPhe 543
QY 2561 GGAAGTAGCCGCTTGTGTAATCTTGCCTTACCTATCGGATCCGATTCGATTAAGGAATCA 2620
DB 544 GlnSerAlaHisPheValAsnValSerLeuProLeuGlyValArgPheGluLysThrAsn 563
QY 2621 GACTGCCAAGATGCAACGTACATCTAACTCTTGGTTTACTGTGGATCTTGTTCGTAGT 2680
DB 564 LysLeuAsn-----ThrTyrAsnIleArgLeuAlaTyrGlnProAspIleTyrArgAsp 581
QY 2681 AACCCGAGCTGTACGACACACTCGCAATTAGCGGTGATTCCTTGAAACCTTCGGTAGC 2740
DB 582 AlaProLysSerLysValPheLeuProSerValHisThrAlaIlePheSerThrGlyAlaThr 601
QY 2741 AATTGGCAAGCAAGCTTTAGTCCTTCGTGCGAGGAACCATTTTTCCTTAACTCAAAT 2800
DB 602 AsnLeuSerArgGlnAlaMetIleLeuAspGlySerAspHisHisLeuThrAspAsn 621
QY 2801 TTTGAAGCCTTTAGCCAAATTTCTTTGAATTCGGTGGTCACTCGCAATTCATATGTA 2860
DB 622 LeuGluValPheCysHisGlyAlaPheGluLeuArgGlySerSerArgAsnTyrAsnVal 641
QY 2861 GACTTAGGACAAAATACCAATTC 2884
DB 642 AspIleGlyArgTyrLysPhe 649
PRELIMINARY; PRT; 186 AA.
```

RESULT 12

Q9RB72

ID Q9RB72

```
AC Q9RB72;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Polymorphic outer membrane protein G family.
GN PMP_3_1.
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RE MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RT Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RL from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314 (2000).
DR EMBL: AP002545; BAA98225.1; -.
DR InterPro: IPR003368; Chlamydia_PMP.
DR Pfam: PF02415; DUF145; 1.
SQ SEQUENCE 186 AA; 19540 MW; 64557A5346533FA3 CRC64;

Alignment Scores:
Pred. No.: 5,68e-46 Length: 186
Score: 917.00 Matches: 185
Percent Similarity: 99.47% Conservative: 1
Best Local Similarity: 98.93% Mismatches: 0
Query Match: 16.95% Indels: 1
DB: 16 Gaps: 0

US-09-428-122-1 (1-3000) x Q9RB72 (1-186)
QY 74 ATTCAAAATCAAGTATATATTTTCAATGAAGTCTTCTTCCCAAGTTGTATTTCT 133
DB 1 MetGlnAsnGlnSerIleTyrPheThrMetLysSerSerPheProLysPheValPheSer 20
QY 134 ACATTTGCTATTTTCCCTTTGCTATGATGCTACCGACAGAGTTTGGATTCAAGTGG 193
DB 21 ThrPheAlaIlePheProLeuSerMetIleAlaThrGluThrValLeuAspSerSerAla 40
QY 194 AGTTTCGATGGGAATAAAAAATGTAATTTTTCAGTTTCGTGAGTCAAGAGATGCTGA 253
DB 41 SerPheAspGlyAsnLysAsnGlyAsnPheSerValArgGluSerGlnGluAspAlaGly 60
QY 254 ACTACTACCTATTTAAGGAAAATGTCCTCTAGAAAATATTCCTGGAAACAGCAGCA 313
DB 61 ThrThrTyrLeuPheLysGlyAsnValThrLeuGluAsnIleProGlyThrGlyThrAla 80
QY 314 ATCACAAAAAGCTGTTTAAACAACACTAAGGCGGATTTGACATTTCACAGGTAAACGGAAC 373
DB 81 IleThrLysSerCysPheAsnAsnThrLysGlyAspLeuThrPheThrGlyAsnGlyAsn 100
QY 374 TCTCTATTGTTCCAAACCGTGGATCGAGGACTGTAGCAGGGGCTGCTTTAACAGCAGC 433
DB 101 SerLeuLeuPheGlnThrValAspAlaGlyThrValAlaGlyAlaAlaValAsnSerSer 120
QY 434 GTGGTAGATAAATCTACACCGTTTATAGGGTTTCTTCGCTATCTTTTATTCGGTCTCT 493
DB 121 ValValAspLysSerThrThrPheIleGlyPheSerSerLeuSerPheIleAlaSerPro 140
QY 494 GGAAGTTCGATTAACACGCAAGAGCGGTTAGTCTGTCTACGGTAGCTTGTAGTTTG 553
DB 141 GlySerSerIleThrThrGlyLysGlyAlaValSerCysSerThrGlySerLeuSerLeu 160
QY 554 ACAAAAAATGTCAGTTTGTCTCTTCAGCAAAAAATCTTCAACGATAATCGCGTGTATC 613
DB 161 Thr-LysMetSerValCysSerSerAlaLysThrPheGlnArgIleMetAlaValLeuSe 180
QY 614 ACCGAAAAATCTTTTCAT 632
DB 180 rProGlnLysLeuPheHis 186
```



```
US-09-428-122-1 (1-3000) x Q9RB68 (1-494)
Qy 164 GCTACCGAGACAGATTTCGATTCAGTCGAGTTTCGATGCGAATAAAATGTAAT--- 220
Db 23 AlaAlaThrThrProLeuAsnProGluAspGlyPheIleGlyGluGlyAsnThrAsnThr 42
Qy 221 TTTTCAGTTCGTGAGAGTCAGAGATGCTGGAACACTACCTACTCTATTTAAAGGGAATGTC 280
Db 43 PheSerProLysSerThrThrAspAlaIleGlyThrThrTySerLeuThrGlyGluVal 62
Qy 281 ACTCTAGAAAATATTCCTGGACACGACAGCAACATCAAAAAGCTGTTTAAACACACT 340
Db 63 ---LeuTyIleAspProGlyLysGlyGlySerIleThrGlyThrCysPheValGluThr 81
Qy 341 AAGGCGGATTGACCTTCACAGTAACGGGACTCTCTATTGTTCCAAACGGTGGATCA 400
Db 82 AlaGlyAspLeuThrPheLeuGlyAsnGlyAsnThrLeuLysPheLeuSerValAspAla 101
Qy 401 GGG---ACTGTAGCAGGGGCTGCTTTACACAGCGGTGGTAGATAAAATCTACCACTTT 457
Db 102 GlyAlaAsnIleAlaValAlaHisValGlnGlySer-----LysAsnLeuSerPhe 118
Qy 458 ATAGGGTTTTCTCGCTAFTCTTTTATGCGTCTCCTGGAACTCGATTAACCGGCAAA 517
Db 119 ThrAspPheLeuSerLeuValIleThrGluSerProLysSerAlaValThrThrGlyLys 138
Qy 518 GGAGCGGTAGTCTCTACGGGTAGCTTTCAGTTTGACAAAATGTCAGTTTGTCTTTC 577
Db 139 GlySerLeu---ValSerLeuGlyAlaValGlnLeuGlnAspIleAsnThrLeuValLeu 157
Qy 578 AGCAAAAATCTTTCACAGGATATGCGGTGCTATCACCGCAAAAATCTTTTCATTAA 637
Db 158 ThrSerAsnAlaSerValGluAspGlyGlyValIleLysGlyAsnSerCysLeuIleGln 177
Qy 638 GGGACTACAATGTCAGCTCTGTTTCTGAAATACCTCTCAAGAAAGCGGAGCACTT 597
Db 178 GlyIleLysAsnSerAlaIlePheGlyGlnAsnThrSerSerLysLysGlyGlyAlaIle 197
Qy 698 CAGACTTCGATCCCTTACCATTTCTGGAACCAAGGGAAGTCTCTTTCTGACAAT 757
Db 198 SerThrThrGlnGlyLeuThrIleGluAsnAsnLeuGlyThrLeuLysPheAsnGluAsn 217
Qy 758 ACTTCTCGGATTCGTGAGTCGAATTTTACAGAAAGCTCGGTGACTATTTCTAATAAT 817
Db 218 LysAlaValThrSerGlyGlyAlaLeuAspLeuGlyAlaAlaSerThrPheThrAlaAsn 237
Qy 818 GCTAAAGTTTCCTTTATGCAATAGGTACAGGAGCGAGCTCC----- 862
Db 238 HisGluLeuIlePheSerGlnAsnLysThrSerGlyAsnAlaAlaAsnGlyGlyAlaIle 257
Qy 863 TCAACACGCGGGATATGTC----- 883
Db 258 AsnCysSerGlyAspLeuThrPheThrAspAsnThrSerLeuLeuLeuGlnGluAsnSer 277
Qy 884 -----GGAGTGTCTATCTGTGCTTATAAAACTAGTACAGATATAAGGTCAAC 931
Db 278 ThrMetClnAspGlyGlyAlaLeuCys-----SerThrGlyThr---IleSer 292
Qy 932 CTCACCTGGAATCAGATGTTACTCTTCAGCAACAATACATCGACAACAGCGGAGGAGCT 991
Db 293 IleThrGlySerAspSerIleAsnValIleGlyAsnThrSerGlyGlnLysGlyGlyAla 312
Qy 992 ATCTATGTGAAAGCTCGAAGTGGCTTCGGA---GGACTTACCTATTTCAGTAGAAT 1048
Db 313 IleSerAlaAlaSerLeuLysIleLeuGlyGlyGlnGlyAlaLeuPheSerAsnAsn 332
Qy 1049 AGTGTCAATGAGGTACAGCTCTTAAAGTGGAGCCATAGCTATCGAAGATAGTGGGAA 1108
Db 333 ValValThrHisAlaThr---ProLeuGlyGlyAlaIlePheIleAsnThrGlyGlySer 351
Qy 1109 TTGAGTTTATCCGCGGATAGTGGTGCATTCCTTTTTCAGGAATACAGTCACTTCTACT 1168
```

```
Db 352 LeuGlnLeuPheThrGlnGlyGlyAspIleValPheGluGlyAsnGlnValThrThr 371
Qy 1169 ACTCTCTGG-----ACGAATAGAAGTAGTATCGACTTAGGAACGAGTGCAGAGATGACA 1222
Db 372 AlaProAsnAlaThrThrLysArgAsnValIleHisLeuGluSerThrAlaLysTrpThr 391
Qy 1223 GCTTTGGTCTGCTGCTGCTAGCCATCTACTCTTCTATGATCCATAACTACAGGATCA 1282
Db 392 GlyLeuAlaAspSerGlnGlyAsnAlaIleThrPheTyAspProIleThrThr---Asn 410
Qy 1283 TCCCAACAGTTACAGATGCTTAAAGTTAAATCAGACTCCGCGAGATTCTCGCACTACAA 1342
Db 411 AspThrGlyAlaSerAspAsnLeuArgIleAsnGluValSerAlaAsnGlnLysLeu--- 429
Qy 1343 TATACAGGAACATCATCTTTCACAGGAGAAAGTATTCACAGACAGAGCGCGAGATTCT 1402
Db 430 ---SerGlySerIleValPheSerGlyGluArgLeuSerThrAlaGluAlaIle---Ala 447
Qy 1403 AAAAATCTTACTTCGAGCTACTACGCTCTTAATCTTTTCAGGAGGTACTCTTCTTTA 1462
Db 448 GluAsnLeuThrSerArgIleAsnGlnProValThrLeuValGluGlySerLeuValLeu 467
Qy 1463 AAACATGAGTGAAGTCTGACAGCTCAGGCATCTCAACAGGAGAGATTCTCGTCTCGAA 1522
Db 468 LysGlnGlyValThrLeuIleThrGlnGlyPheSerGlnGluProGluSerThrLeuLeu 487
Qy 1523 ATGAGCTGAGAACTACTCTA 1543
Db 488 LeuAspLeuGlyThrSerLeu 494
```

RESULT 15

Q9RB70 PRELIMINARY; PRT; 427 AA.

AC Q9RB70; 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Polymorphic outer membrane protein G family.

FM 4_1.

OS Chlamydia pneumoniae (Chlamydophila pneumoniae).

OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.

OX NCBI_TaxID=83558;

RN [1] SEQUENCE FROM N.A.

RC STRAIN=J138;

RX MEDLINE=20330349; PubMed=10871362;

RA Shiba T., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,

RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;

RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138

RL from Japan and CWJ029 from USA.";

RL Nucleic Acids Res. 28:2311-2314 (2000).

DR EMBL; AP002545; BAA98227.1; -

DR InterPro; IPR003368; Chlamydia_PMP.

DR Pfam; PF02415; DUF145; 1.

DR TIGRFAMs; TIGR01376; POMP_repeat; 6.

SQ SEQUENCE 427 AA; 43419 MW; AB4BBBC1594DD2B1 CRC64;

Alignment Scores:

Pred. No.: 1,35e-38 Length: 427

Score: 792.00 Matches: 198

Percent Similarity: 59.56% Conservative: 70

Best Local Similarity: 44.00% Mismatches: 146

Query Match: 14.64% Indels: 36

DB: 16 Gaps: 14

US-09-428-122-1 (1-3000) x Q9RB70 (1-427)

Qy 101 ATGAAGTCTTCTTCCCAAGTTGTATTTTCTACATTTCCTATTTTCCCTTTCTCTATG 160

Db 1 MetArgSerSerPheSerLeuLeuLeuLeuSerSerSerLeuAlaPheProLeuLeuMet 20

Qy 161 ATT-----GCTACCGAGACAGTTTTCGATTTCAGTTCGAGTTCGAGTTCGGAAT 208

Db 21 SerValSerAlaAspAlaAlaAspLeuThrLeuGlySerArgAspSerTyrAsnGlyAsp 40
QY 209 ---AAAATGGTAATTTTCAGTTCGTGAGAGTCAGGAAGATGCT---GGAACTACCTAC 262
Db 41 ThrSerThrThrGluPheThrProLysAlaAlaThrSerAspAlaSerGlyThrThrTyr 60
QY 263 CTATTATAGGGAATGTCACCTAGAAATATCTCTGGAACAGGACACAGCAATACAAAA 322
Db 61 IleLeuAspGlyAspValSerIleSerGln---AlaGlyLysGlnThrSerLeuThrThr 79
QY 323 AGCTGTTTAAACAACACTAAGGCCATTTGACTTTTCACAGGTAAACGGGAATCTCTATTG 382
Db 80 SerCysPheSerAsnThrAlaGlyAsnLeuThrPheLeuGlyAsnGlyPheSerLeuHis 99
QY 383 TTCAAACGGTGGATGCGAGGACTGTAGAGGGGCTGCTGTTTAAACAGCAGCGGTGTAGAT 442
Db 100 PheAspAsnIleIleSerThrValAlaGlyValValSerAsnThrAlaAlaSer 119
QY 443 AAATCTACACGCTTTATAGGTTTCTTCGCTATCTTTTATTCGCTCTCTCGGAAGTTTCG 502
Db 120 GlyIleThrLysPheSerGlyPheSerThrLeuArgMetLeuAlaAlaProArg----- 137
QY 503 ATAACTACCGCAAGAGCCCTAGCTGCTCTACGGGTAGCTTGAGTTTGACAAAAAAT 562
Db 138 ---ThrThrGlyLysGlyAlaIleLysIleThrAspGly---LeuValPheGluSerIle 155
QY 563 GTCAGTTTGCTCTTCAGCAAAACCTTTCAACGGATATATGGGGTGTCTATCACCCAAAA 622
Db 156 GlyAsnLeuAspLeuAsnGluAsnAlaSerSerGluAsnGlyGlyAlaIleAsnThrLys 175
QY 623 ACTCTTTTCATTAAACAGGACTACATGTCAGCTCTGTTTCTGAAATACCTCTCTCAAG 682
Db 176 ThrLeuSerLeuThrGlySerThrArgPheValAlaPheLeuGlyAsnSerSerGln 195
QY 683 AAAGCGGAGCCATTACAGACTTCGATGCCCTTACCATTTACTGTGAAACCAAGGGGAGTC 742
Db 196 GlnGlyGlyAlaIleTyrAlaSerGlyAspSerValIleSerGluAsnAlaGlyIleLeu 215
QY 743 TCTTTTCTGCAATACTTTCGGATTCGTGAGCTGCAATTTTACAGAACCTTCGGTG 802
Db 216 SerPheGlyAsnAsnSerAlaThrThrSerGlyGlyAlaIleSerAlaGluGlyAsnLeu 235
QY 803 ACTATTCTTAATAGCTTAAAGTTTCTTATTTGACATAGGTACAGGACGAGCTCC 862
Db 236 ValIleSerAsnAsnGlnAsnIlePhePheAspGlyCysLysAlaThr----- 251
QY 863 TCAACAACGGGGATATGTCAGGAGTGTCTATCTGTCTTATAAA-----ACTAGTACA 916
Db 252 -----ThrAsnGlyGlyAlaIleAspCysAsnLysAlaGlyAlaAsnPro 266
QY 917 GATCTAAGGTCACTCGGAATCAGATGTTTACTCTTCAGCAACATACATCGACA 976
Db 267 AspProIleLeuThrLeuSerGlyAsnGluSerLeuHisPheLeuAsnAsnThrAlaGly 286
QY 977 ACAGCGGAGGAGCTATCTATGTGAAAAGCTCGAAGTTCGGGA---GGACTTACC 1033
Db 287 AsnSerGlyGlyAlaIleIleThrLysLysLeuValLeuSerSerGlyArgGlyVal 306
QY 1034 CTATTCAGTAGAAATAGTGTCAATGGAGGTACAGCTCTCTAAAGGTGGAGCCATAGCTATC 1093
Db 307 LeuPheSerAsnAsnLysAlaAlaAsnAlaThr---ProLysGlyGlyAlaIleAlaIle 325
QY 1094 GAAGATAGTGGGAATGAGTTTATCCGCCGATAGTGGTGACATGCTCTTTTAGGGAAT 1153
Db 326 LeuAspSerGlyGluIleSerIleSerAlaAspLeuGlyAsnIleIlePheGluGlyAsn 345
QY 1154 ACAGTCACTCTTACTACT-----CCTGGGACCAATAGTAGTAGTATCGACTTAGGA 1204
Db 346 Thr---ThrSerThrThrGlySerProAlaSerValThrArgAsnAlaIleAspLeuAla 364
QY 1205 ACAGTGAAGATGACAGCTTTCGCTGCTGCTGAGAGCCATCTACTTCTATGAT 1264
Db 365 SerAsnAlaLysPheLeuAsnLeuArgAlaThrArgGlyAsnLysValIlePheTyrAsp 384

QY 1265 CCCATTACTACAGATCATCCACACAGTTACAGATGCTTAAAGTTAATGAGACTCCG 1324
Db 385 ProfileThr-----SerSerGlyAlaThrAspLysLeuSerLeuAsnLysAlaAsp 401
QY 1325 GCAGATTCTGCACCTACCAATATACAGGGAACATCATCTTCACAGGAGAAAAGTTATCAGAG 1384
Db 402 AlaGlySerGlyAsnThrTyrGluGlyTyrIleValPheSerGlyGluLysLeuSerGlu 421
QY 1385 ACAGAGGCCGAGATTTCTAAAATCTTACT 1414
Db 422 Val-----ArgAsnLeuThr 426

Search completed: December 16, 2003, 10:27:47

Job time : 168.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2003, 10:33:19 ; Search time 47 seconds
(without alignments)
3134.006 Million cell updates/sec

Title: US-09-428-122-2
Perfect score: 4774
Sequence: 1 MKSFPKVFSTFAIFPLSM.....FELKSSRNYNVDLGAQYQF 928

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :				A_Geneseq_13Jun03.*			
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*				1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*			
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*				2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*			
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*				3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*			
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*				4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*			
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*				5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*			
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*				6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*			
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*				7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*			
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*				8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*			
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*				9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*			
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*				10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*			
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*				11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*			
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*				12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*			
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*				13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*			
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*				14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*			
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*				15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*			
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*				16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*			
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*				17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*			
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*				18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*			
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*				19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*			
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*				20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*			
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*				21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*			
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*				22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*			
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*				23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*			
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*				24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4774	100.0	928	21	AAV94327
2	4760	99.7	928	20	AAW88421
3	3865	81.0	746	23	ABB90535
4	2802.5	48.1	597	20	AAV34611
5	2058	43.1	928	20	AAW88418
6	2048	42.9	928	23	ABB90573
7	2021	42.3	928	21	AAV90237
8	2006	42.0	918	21	AAV69369
9	2001	41.9	918	20	AAW88422

10	1993	41.7	914	20	AAW88429	Chlamydia pneumoni
11	1987.5	41.6	885	21	AAV90238	Mature Chlamydia a
12	1986	41.6	928	20	AAW88423	Chlamydia pneumoni
13	1982	41.5	928	23	ABB90542	Chlamydia pneumoni
14	1976	41.4	928	21	AAV90239	Chlamydia antigen
15	1965	41.2	936	21	AAV99842	Chlamydia pneumoni
16	1965	41.2	936	23	ABB90602	Chlamydia pneumoni
17	1946.5	40.8	925	21	AAV99843	Chlamydia pneumoni
18	1940	40.6	930	20	AAV35052	Chlamydia pneumoni
19	1936	40.6	930	23	ABB90548	Chlamydia pneumoni
20	1930	40.4	930	21	AAV90240	Chlamydia antigen
21	1927	40.4	930	20	AAW88424	Chlamydia pneumoni
22	1917.5	40.2	927	20	AAV35054	Chlamydia pneumoni
23	1915	40.1	926	23	ABP56019	Chlamydia psittaci
24	1915	40.1	926	23	ABB98228	Chlamydia polypept
25	1915	40.1	926	24	ABU66284	C. psittaci protei
26	1856	38.9	949	20	AAV35060	Chlamydia pneumoni
27	1855	38.9	928	20	AAW88417	Chlamydia antigen
28	1855	38.9	928	21	AAV90236	Chlamydia pneumoni
29	1855	38.9	928	23	ABB90583	Amino acid sequenc
30	1832	38.4	945	21	AAV69368	Chlamydia pneumoni
31	1811	37.9	945	20	AAW88428	Chlamydia psittaci
32	1656.5	34.7	839	23	ABP56002	Chlamydia polypept
33	1656.5	34.7	839	23	ABB98211	C. psittaci protei
34	1656.5	34.7	839	24	ABU66267	Chlamydia pneumoni
35	1573	32.9	841	23	ABB90595	Chlamydia pneumoni
36	1572.5	32.9	643	20	AAV35056	Chlamydia pneumoni
37	1570	32.9	841	21	AAV92818	C. pneumoniae CPN1
38	1564	32.8	841	20	AAW88420	Chlamydia pneumoni
39	1442.5	30.2	922	23	ABB90546	Chlamydia pneumoni
40	1437.5	30.1	922	21	AAV95548	Chlamydia pneumoni
41	1430.5	30.0	922	20	AAW88419	Chlamydia pneumoni
42	1417.5	29.7	922	20	AAV34597	Chlamydia pneumoni
43	1383.5	29.0	1407	23	ABB90541	Chlamydia pneumoni
44	1377.5	28.9	973	21	AAV96274	Chlamydia POMP91B
45	1377.5	28.9	973	23	ABB90527	Chlamydia pneumoni

ALIGNMENTS

RESULT 1
AAV94327
ID AAV94327 standard; Protein; 928 AA.
XX AAV94327;
AC
XX
DT 11-AUG-2000 (first entry)
XX
DE Chlamydia pneumoniae 98kD putative outer membrane protein.
XX
KW Chlamydia; antigen; vaccine; infection; outer membrane protein.
XX
OS Chlamydia pneumoniae.
XX
PN WO200026237-A2.
XX
PD 11-MAY-2000.
XX
PF 29-OCT-1999; 99WO-GB03579.
XX
PR 29-OCT-1998; 98US-0106070.
PR 01-MAR-1999; 99US-0122066.
PR 27-OCT-1999; 99US-0428122.
XX
(CONN-) CONNAUGHT LAB LTD.
XX
PI Murdin AD, Oomen RP, Dunn PL;
XX
XX WPI; 2000-365569/31.
DR N-PSDB; AAA27021.
XX
PT Novel Chlamydia 98 kDa putative outer membrane protein antigen, used

for vaccination and protection against Chlamydia infection

Claim 6; Fig 1; 93pp; English.

The present sequence is the 98kDa putative outer membrane protein from Chlamydia pneumoniae. The genomic sequence was amplified using two PCR primers. The 5' primer contains a NotI restriction site, a ribosome binding site, an initiation codon and a sequence close to the 5' end of the 98kDa putative outer membrane protein coding sequence. The 3' primer contains the sequence encoding the C-terminal sequence of the putative outer membrane protein and a BglII restriction site. The stop codon was excluded and an additional nucleotide was inserted to obtain an in-frame C-terminal fusion with the Histidine tag. The PCR product was cloned into a eukaryotic expression vector (pCA-Myc-His) by restricting both the vector and the PCR product with NotI and BamHI and performing a ligation reaction. This expression vector was injected intramuscularly and intranasally into mice, which were subsequently inoculated with Chlamydia pneumoniae. The chlamydial lung titers of the immunised mice were lower than those of the controls. Thus the 98kDa putative outer membrane protein can be used as a vaccine to provide protection against Chlamydia infections, especially Chlamydia pneumoniae infections. The present polypeptide may also be administered orally to treat Chlamydia infection.

Sequence 928 AA;

Query Match 100.0%; Score 4774; DB 21; Length 928;

Best Local Similarity 100.0%; Pred. No. 6.5e-305;

Matches 928; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MKSSFPKVFSTFAIPPLSMIATETVLDSSAGFDGKNGNFSVRESQEDAGTYLPKGNV	60
Db	1	MKSSFPKVFSTFAIPPLSMIATETVLDSSAGFDGKNGNFSVRESQEDAGTYLPKGNV	60
QY	61	TLENIPGTGTATKSCFNNTKGDLTFTGNSLLFQTVAGTAVAGAVNSVVDKSTTFI	120
Db	61	TLENIPGTGTATKSCFNNTKGDLTFTGNSLLFQTVAGTAVAGAVNSVVDKSTTFI	120
QY	121	GRSSLSFIAPSGSIITGKGVSCSTGSLSTKNVLLFSKNFSTDPNGGAIKTAKTLSLGT	180
Db	121	GRSSLSFIAPSGSIITGKGVSCSTGSLSTKNVLLFSKNFSTDPNGGAIKTAKTLSLGT	180
QY	181	TTMSALFSENTSSKKGAIQTSDALITGNGEVSPDNTSSDGGAAITFEASVTISNNA	240
Db	181	TTMSALFSENTSSKKGAIQTSDALITGNGEVSPDNTSSDGGAAITFEASVTISNNA	240
QY	241	KVSFIDNKVTGASSSTTGDMSGGAIKAYKTSTDTKVTLTGNQMLLFSNNTSTTAGGAIYV	300
Db	241	KVSFIDNKVTGASSSTTGDMSGGAIKAYKTSTDTKVTLTGNQMLLFSNNTSTTAGGAIYV	300
QY	301	KKLELASGGLTFSRNSVNGTAPKGGAIKAIEDSGELSLSDSGDIVFLGNTVSTTPGT	360
Db	301	KKLELASGGLTFSRNSVNGTAPKGGAIKAIEDSGELSLSDSGDIVFLGNTVSTTPGT	360
QY	361	NRSSIDLGTSKMTALRSAGRAIYFVDPITTGSSITVTVLKNETPADSALOYTGNI	420
Db	361	NRSSIDLGTSKMTALRSAGRAIYFVDPITTGSSITVTVLKNETPADSALOYTGNI	420
QY	421	FTGEKLSSETAADSCKNLTSLKLOPVLTSGGTSLKHGVTIQTQFTQOADSRLMDVGT	480
Db	421	FTGEKLSSETAADSCKNLTSLKLOPVLTSGGTSLKHGVTIQTQFTQOADSRLMDVGT	480
QY	481	LEPADTGTINNLVINISIDGAKKAKIETKATSKNLTLSGTITILLDPTGTFYENHSRLNP	540
Db	481	LEPADTGTINNLVINISIDGAKKAKIETKATSKNLTLSGTITILLDPTGTFYENHSRLNP	540
QY	541	QSYDILELKASGTVTSTAVTDPDPIMGEKPHYGYGTGPIVWGTGASTTATFNWTKGYI	600
Db	541	QSYDILELKASGTVTSTAVTDPDPIMGEKPHYGYGTGPIVWGTGASTTATFNWTKGYI	600
QY	601	PNPERIGSLVNSLWNAFIDISLHYLMETANEGLOGDRAFCWAGLNFHFKDSTKTRRG	660
Db	601	PNPERIGSLVNSLWNAFIDISLHYLMETANEGLOGDRAFCWAGLNFHFKDSTKTRRG	660

QY	661	FRHLSGGYVIGGNLHTCSDKILSAFCQLFGRDRDYFVAKNQGTGYGGTLYYQHNETYIS	720
Db	661	FRHLSGGYVIGGNLHTCSDKILSAFCQLFGRDRDYFVAKNQGTGYGGTLYYQHNETYIS	720
QY	721	LPCKLRPCSLSYVPTFPLVPSGNLSYTHTDNDLTKYTYTPTVKGSGWGNDSFALEFGGR	780
Db	721	LPCKLRPCSLSYVPTFPLVPSGNLSYTHTDNDLTKYTYTPTVKGSGWGNDSFALEFGGR	780
QY	781	APICLDESALFEQYMPFMKIQFYVAHQEGFKQTEAREFGSSRLVNLALPIGIRFPDKES	840
Db	781	APICLDESALFEQYMPFMKIQFYVAHQEGFKQTEAREFGSSRLVNLALPIGIRFPDKES	840
QY	841	DCQDATYNLTGYTVDLVRSNPDCTTTLRISGDSWKTEGTNLARQALVLRAGNHPCFNSN	900
Db	841	DCQDATYNLTGYTVDLVRSNPDCTTTLRISGDSWKTEGTNLARQALVLRAGNHPCFNSN	900
QY	901	FEAFSQSFELRGSSRNYNVDLGAQYQF	928
Db	901	FEAFSQSFELRGSSRNYNVDLGAQYQF	928

RESULT 2

AAW88421

ID AAW88421 standard; Protein; 928 AA.

XX AAW88421;

XX 26-APR-1999 (first entry)

XX Chlamydia pneumoniae surface exposed protein Omp8.

XX Omp8; outer membrane protein 8; surface exposed protein; antigen; infection; diagnosis; vaccine; atherosclerosis; asthma.

XX Chlamydia pneumoniae.

XX WO9858953-A2.

XX 30-DEC-1998.

XX 19-JUN-1998; 98WO-DK00266.

XX 23-JUN-1997; 97DK-0000744.

XX (BIRK/) BIRKELUND S.

XX (CHRI/) CHRISTIANSEN G.

XX Birkelund S, Christiansen G, Knudsen K, Madsen A;

XX Mygind P;

XX WPI; 1999-105610/09.

XX N-PSDB; AAX06820.

XX Species-specific test for identifying mammals infected with Chlamydia pneumoniae - comprises detecting antibodies specific for outer membrane proteins of C. pneumoniae or nucleic acids encoding these proteins

XX Claim 7; Page 53-55; 115pp; English.

XX This polypeptide comprises the novel 90.0 kDa surface exposed protein Omp8 of the human respiratory pathogen Chlamydia pneumoniae. Its amino acid sequence was deduced from DNA (see AAX06820) isolated from a C. pneumoniae expression library. The invention provides 12 novel surface exposed proteins, Omp4-Omp15 (see AAW88417-28), and nucleic acid sequences encoding them (see AAX06816-27). A new species specific test is claimed that is used to identify mammals (including humans) infected with Chlamydia pneumoniae. The test comprises detecting antibodies specific for Omp4-Omp15 or detecting nucleic acid fragments encoding these outer membrane proteins, especially by PCR. The proteins are also used in the diagnosis of C. pneumoniae infection in mammals. The

CC nucleic acids and proteins can also be used in the immunization of
CC mammals, the nucleic acids being particularly useful as DNA
CC vaccines for effecting *in vivo* expression of antigens. The
CC vaccines may also prevent atherosclerosis and bronchial asthma,
CC which are possibly associated with *C. pneumoniae*.

Sequence	928 AA;
SO	

Query Match	99.7%;	Score 4760;	DB 20;	Length 928;
Best Local Similarity	99.7%;	Pred. No. 5.4e-304;		
Matches 925; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

Qy	1	MKSFPKVFSTFAIFPLSMIATETVLDSSAFDNKNKGNFVSRESQEDAGTTVLFKGNV	60
Db	1	MKSFPKVFSTFAIFPLSMIATETVLDSSAFDNKNKGNFVSRESQEDAGTTVLFKGNV	60
Qy	61	TLENIPTGTATKSCFNNTKGDLTFTGNGNSLLFQTVDACTVAGAAVSSVDDKSTTFI	120
Db	61	TLENIPTGTATKSCFNNTKGDLTFTGNGNSLLFQTVDACTVAGAAVSSVDDKSTTFI	120
Qy	121	GFSSLSFTASPGSSTTTCKGAVSCSTGSLSTKNVSLLFKNFSFDNGGAIATAKTLSLTG	180
Db	121	GFSSLSFTASPGSSTTTCKGAVSCSTGSLKPKDNVSLLFKNFSFDNGGAIATAKTLSLTG	180
Qy	181	TTMSALFSGENTSSKKGAIQTSDALTTTGNQGEVFSFSDNTSSDSGAAIIFTEASVTISNNA	240
Db	181	TTMSALFSGENTSSKKGAIQTSDALTTTGNQGEVFSFSDNTSSDSGAAIIFTEASVTISNNA	240
Qy	241	KVSFTDNKVTGASSSTTGDMSGGAIKAYKTSFDTKVTLTGNQMLLFNSNTSTTAGGAIYV	300
Db	241	KVSFTDNKVTGASSSTTGDMSGGAIKAYKTSFDTKVTLTGNQMLLFNSNTSTTAGGAIYV	300
Qy	301	KKLEIASGGLTLFNRNSVNGGTPAKGGAIAIEDSGELSLSDSGDIFVLGNTVTTSTPGT	360
Db	301	KKLEIASGGLTLFNRNSVNGGTPAKGGAIAIEDSGELSLSDSGDIFVLGNTVTTSTPGT	360
Qy	361	NRSSIDLQPSAKWTALRSAAGRAIYPYDPIITGSSSTTVTDVLKVNETPADSALQVTGNI	420
Db	361	NRSSIDLQPSAKWTALRSAAGRAIYPYDPIITGSSSTTVTDVLKVNETPADSALQVTGNI	420
Qy	421	FTGEKLSFTEAADSKNLTSLKLPVTLSGGTLSLKHGVTLQATFQQQADSLFMDVGT	480
Db	421	FTGEKLSFTEAADSKNLTSLKLPVTLSGGTLSLKHGVTLQATFQQQADSLFMDVGT	480
Qy	481	LEPADTSTNNLVINISSIDGAKKAKIETKATSKNLTLSGTTLLDPTGTFYENHSLRNP	540
Db	481	LEPADTSTNNLVINISSIDGAKKAKIETKATSKNLTLSGTTLLDPTGTFYENHSLRNP	540
Qy	541	QSYDILLELKASGTVTSTAVIDPIMGEKPHYGQGTWGPVWGTVGASTTATFNWTKGYI	600
Db	541	QSYDILLELKASGTVTSTAVIDPIMGEKPHYGQGTWGPVWGTVGASTTATFNWTKGYI	600
Qy	601	PNPERIGSLVPNSLWNAFIDISLHYLMETANGLQDRAFWCAGLSNPFHKDSTKTRRG	660
Db	601	PNPERIGSLVPNSLWNAFIDISLHYLMETANGLQDRAFWCAGLSNPFHKDSTKTRRG	660
Qy	661	FRHLSGGYVIGGNLHTCSDKILSAAFCQLFGRDRDYFVAKNQGTVYGGTLYYQHNFTYIS	720
Db	661	FRHLSGGYVIGGNLHTCSDKILSAAFCQLFGRDRDYFVAKNQGTVYGGTLYYQHNFTYIS	720
Qy	721	LPCKLRPCSLSYVPTPIVLFSGNLSYTHTDNDLKTXYTTPYVKSGWGNDSFALFSGGR	780
Db	721	LPCKLRPCSLSYVPTPIVLFSGNLSYTHTDNDLKTXYTTPYVKSGWGNDSFALFSGGR	780
Qy	781	APICLDSALFEQWMPFMKQFVYAHQGEFKQTEAREFGSSRLVNIALPIGIRDPKES	840
Db	781	APICLDSALFEQWMPFMKQFVYAHQGEFKQTEAREFGSSRLVNIALPIGIRDPKES	840
Qy	841	DCQDATYNLTLYGTVDLVRNPDCTTTILRISGDSWKTFGTNLARQALVLRAGNHFCSNS	900
Db	841	DCQDATYNLTLYGTVDLVRNPDCTTTILRISGDSWKTFGTNLARQALVLRAGNHFCSNS	900
Qy	901	FEAFSQFSFELRGSSRNYNVDLGAKYQF	928

db 901 FEAFSQFSFELRGSSRNYNVDLGAKYQF 928

RESULT 3
ABB90535

ABB90333
ID ABB90535 standard; Protein; 746 AA.

AC ABB90535;

29-III.-2002 (first entry)

Chlamydia pneumoniae cp0015 protein, SEQ ID NO:19.

XX Chlamydial infection; antigen; immunogen; vaccine; diagnosis;
KW human respiratory disease; cardiovascular disease; atherosclerosis;
KW coronary artery disease; carotid artery stenosis; myocardial infarction;
KW cerebrovascular disease; aortic aneurysm; claudication; stroke;
KW strain CWL029.

OS Chlamydia pneumoniae.

PN WO200202606-A2.

PD 10-JAN-2002.

03-JUL-2001: 2001WO-IB01445.

03-ТІІІ.-2000: 2000GB-0016363.

PR II-JUL-2000; 2000GB=0017983
PR 31-JUL-2000: 2000GB=0017983

PR 07-AUG-2000; Z000GB-00I3388
PB 18-AUG-2000: Z000GB-0020440

PR I4=SELF=2000; Z000GB=0022383
DB 10-NOV-2000: 3000GB=0027549

PR 22-DEC-2000; 2000GB-0031708-
vv

PA (CHIR-) CHIRON SPA.

PI Ratti G, Grandi G;
1997

DR WPI; 2002-154726/20.

XXIX

PT medicament for treatment or prevention of infection due to Chlamydia pneumoniae, and for diagnostic purposes -

PS Claim 1; Page 49-50; 364pp; English.

Sequences ABB90526-AB890715 represent novel proteins from *Chlamydia pneumoniae* (strain CWL029), and ABL91184-ABL91373 represent *Chlamydia pneumoniae*. The proteins are predicted to be immunogenic and may therefore be useful in vaccine production and for diagnostic purposes. *Chlamydia pneumoniae* is a common cause of respiratory disease in humans, and is also involved in the development of cardiovascular diseases such as atherosclerosis, coronary artery disease, carotid artery stenosis, myocardial infarction, cerebrovascular disease, aortic aneurysm, claudication and stroke. The proteins and nucleic acids of the invention may be used in vaccines and pharmaceutical compositions for the prevention or treatment of *Chlamydia* infections, particularly *Chlamydia pneumoniae* infections. The proteins may also be used in the detection of *Chlamydia pneumoniae*, and the nucleic acids may be used in PCR, branched DNA probe assay or blotting techniques for determining *Chlamydia pneumoniae* gene expression. The present sequence represents a specifically claimed *Chlamydia pneumoniae* protein of the invention.

AA	Sequence	746 AA:
SO		

Query Match 81.0%: score 3865: DB 23: Length 746:

Best Local Similarity 100.0%; Freq. NO: 2:36-2437
Matches 746; Conservative 0; Mismatches 0:

901 FEAFSOFSELRGSSRNYNVDLGAKYQF 928 Ov

183	QY	MSALFSENTSSKKGAIQTS	DALPIITGNQGEVSFSDNTSS	DGAAIFTEASVTISNNKV	242			
	Db	1	MSALFSENTSSKKGAIQTS	DALPIITGNQGEVSFSDNTSS	DGAAIFTEASVTISNNKV	60		
243	QY	SFIDNKVTGASSSTGDM	SGGAI	CAYKTSTDKVTLTGNQMLLFS	NNSTSTAGGAIYVKK	302		
	Db	61	SFIDNKVTGASSSTGDM	SGGAI	CAYKTSTDKVTLTGNQMLLFS	NNSTSTAGGAIYVKK	120	
303	QY	LELASGGTLFSRNVNG	GTAPKGAIAIEDSGELSL	SADSGDIVFLGNTVTS	TPGTNR	362		
	Db	121	LELASGGTLFSRNVNG	GTAPKGAIAIEDSGELSL	SADSGDIVFLGNTVTS	TPGTNR	180	
363	QY	SSIDLGTSAKMTAL	NSAAGRAIFYDPIITGSS	TTVDVLKVNETPAD	SALQYTGNIIFT	422		
	Db	181	SSIDLGTSAKMTAL	NSAAGRAIFYDPIITGSS	TTVDVLKVNETPAD	SALQYTGNIIFT	240	
423	QY	GKLISETAADS	KNLTSKLQPVTLISGGT	LSLKHGVTLQTAFTQOADS	RLMDVGTILE	482		
	Db	241	GKLISETAADS	KNLTSKLQPVTLISGGT	LSLKHGVTLQTAFTQOADS	RLMDVGTILE	300	
483	QY	PADTSTINNLVINI	SSIDGAKKAKIETKAT	SKNLTLSGTITLLDPTG	TFYENHSLRNPOS	542		
	Db	301	PADTSTINNLVINI	SSIDGAKKAKIETKAT	SKNLTLSGTITLLDPTG	TFYENHSLRNPOS	360	
543	QY	YDILELKASG	TVTSTA	VPDPIMGSKFHYG	GTWGPVWGTGAST	TATFNTKTYIPN	602	
	Db	361	YDILELKASG	TVTSTA	VPDPIMGSKFHYG	GTWGPVWGTGAST	TATFNTKTYIPN	420
603	QY	PERIGSLVPSI	LWNAFIDISSLHYLMET	ANEGLQGDRAFCWAGLS	FNFFHKDSTKTRGR	662		
	Db	421	PERIGSLVPSI	LWNAFIDISSLHYLMET	ANEGLQGDRAFCWAGLS	FNFFHKDSTKTRGR	480	
663	QY	HLSGGYVIGGNL	HTCSDKILSAACQL	FCRDRDYFVAKNQGT	VYGGTLYYQHNETYISLP	722		
	Db	481	HLSGGYVIGGNL	HTCSDKILSAACQL	FCRDRDYFVAKNQGT	VYGGTLYYQHNETYISLP	540	
723	QY	CKLRPCSL	SVYPTIEIVLFSGNLS	YTHTDNLTKKTYTTPY	VKSGMGNDSFALEFGGRAP	782		
	Db	541	CKLRPCSL	SVYPTIEIVLFSGNLS	YTHTDNLTKKTYTTPY	VKSGMGNDSFALEFGGRAP	600	
783	QY	ICLDESAL	FEQYMPFMKLOFVY	AHOBGFKEQGT	EAREFGSSRLVNLALPIGIR	FPKESDC	842	
	Db	601	ICLDESAL	FEQYMPFMKLOFVY	AHOBGFKEQGT	EAREFGSSRLVNLALPIGIR	FPKESDC	660
843	QY	QDATYNLT	LGTYVDLVRSPDC	TTTLIRISGDS	WKTEGTNLARQALVLRAGNH	FCFNSNFE	902	
	Db	661	QDATYNLT	LGTYVDLVRSPDC	TTTLIRISGDS	WKTEGTNLARQALVLRAGNH	FCFNSNFE	720
903	QY	AFSQFSE	FELRGSSRN	NVDLGAKYQF	928			
	Db	721	AFSQFSE	FELRGSSRN	NVDLGAKYQF	746		

RESULT 4	
AAAY34611	standard; Protein; 597 AA.
ID	AAAY34611
XX	
XX	AAAY34611;
XX	
XX	AC
XX	AC
DT	13-SEP-1999 (first entry)
XX	
XX	Chlamydia pneumoniae transmembrane protein sequence.
DE	
XX	
XX	Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW	sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW	vaccine; neutralising epitope.
KW	
XX	
OS	Chlamydia pneumoniae.
OS	
XX	
PN	W09927105-A2.
XX	
XX	
PD	03-JUN-1999.
XX	
XX	

PF	20-NOV-1998;	98WO-IB01890.
XX		
PR	04-NOV-1998;	98US-0107078.
PR	21-NOV-1997;	97FR-0014673.
XX		
PA	(GEST) GENSET.	
XX		
PI	Griffais R;	
XX		
DR	WPT; 1999-357842/30.	
XX		
PT	Genome sequence of Chlamydia pneumoniae	
XX		
PS	Page 632-633; Disclosure; 1912pp; English.	
XX		
CC	AAV34584-Y35879 represent the proteins encoded by all the open reading	
CC	frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.	
CC	C. pneumoniae causes respiratory disease such as pneumonia and	
CC	bronchitis and is thought to be a contributing factor in heart	
CC	disease, sarcoidosis, sinusitis, purulent otitis media, erythema	
CC	nodosum or pharyngitis. The polypeptides encoded by the open reading	
CC	frames of the C. pneumoniae genome (see AAX34584-Y35879) can be used in	
CC	immunogenic compositions as vaccines. Vectors containing C. pneumoniae	
CC	nucleotide sequences can also be used as immunogenic compositions,	
CC	especially where the vector directs the expression of a neutralising	
CC	epitope of C. pneumoniae.	
XX		
SQ	Sequence 597 AA;	
XX		
<hr/>		
	Query Match	Score 2802.5; DB 20; Length 597;
	Best Local Similarity	93.8%; Pred. No. 1.2e-175;
	Matches 560:	Conservative 5; Mismatches 21; Indels 11; Gaps 3;

QY	149	LELTQVSLLSKPNFTDNGGAI	TA	KTLSLTGTTWMSALFSEN	TS	SKKGGAIQTSDAL	TT	208
DB	1	LEFDKNVSLLSKPNFTDNGGAI	TA	KTLSLTGTTWMSALFSEN	TS	SKKGGAIQTSDAL	TT	60
QY	209	GNQGEVSFSDNTSSDSGAAI	FTEASVT	ISNNAKVSFIDNKVT	GASSST	IGDMSGGAICAY		268
DB	61	GNQGEVSFSDNTSSDSGAAI	FTEASVT	ISNNAKVSFIDNKVT	GASSST	IGDMSGGAICAY		120
QY	269	KTSTDTKVTLTGQMLFSSNT	STTAGGAI	VYKLELASGG	TLF	SRNSVNGGTA	PKGGA	328
DB	121	KTSTDTKVTLTGQMLFSSNT	STTAGGAI	VYKLELASGG	TLF	SRNSVNGGTA	PKGGA	180
QY	329	IAIEDSGELSLADSGDIV	FLGNTV	STTPGNNRSIDLGT	SAKMTAL	RSAGRAI	VEYD	388
DB	181	IAIEDSGELSLADSGDIV	FLGNTV	STTPGNNRSIDLGT	SAKMTAL	RSAGRAI	VEYD	240
QY	389	PIITGSGTIVTDVLKVN	ETPADSALQY	TGNIIFTGEKL	SETEAADSKNLT	SKLQP	VTLS	448
DB	241	PIITGSGTIVTDVLKVN	ETPADSALQY	TGNIIFTGEKL	SETEAADSKNLT	SKLQP	VTLS	300
QY	449	GGTLSLKHGVTLTQTAFT	QOADSRL	EMDVGTTLEPAD	TSTINNLVIN	ISSIDGAKAKIE		508
DB	301	GGTLSLKHGVTLTQTAFT	QOADSRL	EMDVGTTLEPAD	TSTINNLVIN	ISSIDGAKAKIE		360
QY	509	TKATSKNLTLSGTITILL	DPGTGFYENHSLRN	POSVDILLEKASG	VTSTAVTPDP	IMGEK		568
DB	361	TKATSKNLTLSGTITILL	DPGTGFYENHSLRN	POSVDILLEKASG	VTSTAVTPDP	IMGEK		420
QY	569	FHYGYOGTWGPIVWGT	GASTATFNWTKGT	YIPNPERIGSLVPNS	LWNAFIDIS	SLHYLM		628
DB	421	FHYGYOGTWGPIVWGT	GASTATFNWTKGT	YIPNPERIGSLVPNS	LWNAFIDIS	SLHYLM		480
QY	629	ETANEGLQGDRAFCW	AGLSNFPHKOSTKTR	RGRFHL	SGGYVIGGNL	HTCSDKIL	SAAPCO	688
DB	481	ETANEGLQGDRAFCW	AGLSNFPHKOSTKTR	RGRFHL	SGGYVIGGNL	HTCSDKIL	SAAPCO	540
QY	689	LFGRDRDYFVAKNQG	-TVVYGGTILY	QHNETHLS-----LPCKLR	-PCSLSYVP			734
DB	541	LFGRDRDYFVAKNQY	SLRNSLSPALRN	LYLSLQTTALFV	LVLCYRSDCS	LFKPK		597

RESULT 5

AAW88418
ID AAW88418 standard; Protein; 928 AA.
XX
AC AAW88418;
XX
DT 26-APR-1999 (first entry)
XX
DE Chlamydia pneumoniae surface exposed protein Omp5.
XX
DE Omp5; outer membrane protein 5; surface exposed protein; antigen;
KW infection; diagnosis; vaccine; atherosclerosis; asthma.
XX
KW Chlamydia pneumoniae.
OS
XX
XX WO9858953-A2.
XX
XX 30-DEC-1998.
PD
XX
XX 19-JUN-1998; 98WO-DK00266.
PF
XX
XX 23-JUN-1997; 97DK-0000744.
PR
XX
XX (BIRK/) BIRKELUND S.
PA (CHRI/) CHRISTIANSEN G.
XX
XX Birkelund S, Christiansen G, Knudsen K, Madsen A;
PI Mygind P;
XX
XX WPI: 1999-105610/09.
DR N-PSDB; AAX06817.
XX
XX
XX Species-specific test for identifying mammals infected with
PT Chlamydia pneumoniae - comprises detecting antibodies specific for
PT outer membrane proteins of C. pneumoniae or nucleic acids encoding
PT these proteins
XX
PS Claim 7; Page 43-45; 115pp; English.
XX
XX This polypeptide comprises the novel 97.2 kDa surface exposed
CC protein Omp5 of the human respiratory pathogen Chlamydia
CC pneumoniae. Its amino acid sequence was deduced from DNA (see
CC AAX06817) isolated from a C. pneumoniae expression library. The
CC invention provides 12 novel surface exposed proteins, Omp4-Omp15
CC (see AAW88417-28), and nucleic acid sequences encoding them (see
CC AAX06816-27). A new species specific test is claimed that is used
CC to identify mammals (including humans) infected with Chlamydia
CC pneumoniae. The test comprises detecting antibodies specific for
CC Omp4-Omp15 or detecting nucleic acid fragments encoding these outer
CC membrane proteins, especially by PCR. The proteins are also used
CC in the diagnosis of C. pneumoniae infection in mammals. The
CC nucleic acids and proteins can also be used in the immunization of
CC mammals, the nucleic acids being particularly useful as DNA
CC vaccines for effecting in vivo expression of antigens. The
CC vaccines may also prevent atherosclerosis and bronchial asthma,
CC which are possibly associated with C. pneumoniae.
XX
XX Sequence 928 AA;
XX
Query Match 43.1%; Score 2059; DB 20; Length 928;
Best Local Similarity 46.7%; Pred. No. 1.9e-126;
Matches 444; Conservative 142; Mismatches 320; Indels 44; Gaps 20;
XX
QY 1 MKSFFPKFVF-STFAIF--PLSMIATETVLDSSASFDGNKN-GNFSVRESQBDAGTYL 55
DB 1 MKQFSLVLSLTLACFTSCSTVFAAENIGFSDSGTNGTYTPKNT--TTGIDYT 58
QY 56 FKGNTVLENIPTGTALTATKCFNNKGDLLFTNGNLSLLFTQVDAGTVAGAAVNSVVDK 115
DB 59 LTGDTITQNL-GDSAAITKGCFSDDTESLSFAGKGYSLFLINKS-SAEGAAL-SVTTDK 115
QY 116 STTFIGSSLSFIASPOSSITT--GKAVSCSTGSLSLTKNLSLLFSKNFSTDNGGAI 173

Db 116 NLSLTGFSSLTFLAAPSVTTPSGKAVKCG-GDLTFDNNGTILFKQDYCEENGGAIST 174
QY 174 KTLISLTGTTWTSALPSENTSS---KKGGAIQTSALDALTITGNGEVSFSDNTSSDGAIFT 230
Db 175 KNLSLKNSTGSISEFEGNKSATGKKGAICATGTVDTNNAPTFLFNNIAEAGGAINS 234
QY 231 EASVTISNNAKVSFIDNKVTGASSSTTGDMSGGAI CAYKTSTDTKVTLTGQMLLFSNNT 290
Db 235 TNCITITGNTSLVSENSVT---ATAG--NGGAL-----SGDADVIISNQSVTFSGNQ 283
QY 291 STTAGGAIYVKLELAS--GGLTLFNRNSVNGTAPKGAIAIEDSGELSASAGSDIVF 348
Db 284 AVANGGAIYAKKLTLAGGGGGGGSFNINVGTTAGNGGALSILAAGBCSLAAGADITF 343
QY 349 LGNTVTTSTTP-GTNRSSIDLGTSAKMTALRAAGRAIFYDYDPIITGSSSTTVDVLKNET 407
Db 344 NGNAIVATTPQTTRKNSIDIGSTAKITNLRAISGHSIFFYDPIITANTAADSTDTLNLNKA 403
QY 408 PADSALOYTGNIIFTGEKLSSETAADSKNLTSKLLQPVTLSSGTLISLKHGVTLTQAFATQ 467
Db 404 DAGNSTDYGSIYVSGEKLSEDAKVADNLITSLKQPVTLTAGNLVLRGVTLDTKGTQ 463
QY 468 QADSRLEMDVGTTLLEPA-DTSTINNLVINISIDGAKKAKIETKATSKNLTLSGHTILLD 526
Db 464 TAGSSVIMDAGTTLKASTEEVTLGLSLPVDLSLGEKKVIAAASAASKNVALSGPILLD 523
QY 527 PTGTFYENHSLRNQSYDILELKASGTVTSTAVTPDPIMGEKHFHYGYOGTWGPVW---- 582
Db 524 NQGNAYENHDLGKTQDFSVQLSALGTATTDVPAVPTVATPTHYGYOGTWG-MTWVDDT 582
QY 583 -GTGASTTATFNWTKTGYIPNPERIGSLVPSNLSMNAFIDISLHYLMETANEGLQDRAF 641
Db 583 ASTPKTKATLAWTNGVLPNPERQGPLVPSNLSMGSPSDIQAIQGVIERALSALTCLSDRGF 642
QY 642 WCAGLSNPFHKDSTKTRGRFHLSGGYVIGNMLHTCSDKILSAAPCOLFGRDRDYFAKN 701
Db 643 WAAGVANFLDKKGKKRYRHKSGGYAIGAAQTCSENLISFAFCQLFGSKDFLVAKN 702
QY 702 QGTVVGTYLYOHNETYISLPCKLRPCSLSYVP---TEIPVLFSGNLSYTHTDNDLKTXY 758
Db 703 HTDTYAGAFYIQH---ITECSGFGCLDLKPLGWSHKLPLVEQLAYSHVSNDLTKY 758
QY 759 TTYPTVKSGWGNDSFALBFGGRAPICLDBESALFEQVMPFMKLFQVYAHGEFKEQTEAR 818
Db 759 TAYPEVKSNGWNAFNMMLGASHSYPEYLHCFDIYAPYIKLNLTYIRODSFSEKGTGR 818
QY 819 EFGSRLVNLALPIGIRFDKESDQDQATYNLTGTYVDLVLRNPDCTTTLRISGDSWKTF 878
Db 819 SFDDSNLFLNLSLPIGVKFEKPSDCNDFSVDLTLSYVPDLIRNDPKCTTALVISGASWET 878
QY 879 GTNLARQALVLRAGNHCFNSNFEAFSPFELRGSSNNVVDLGAKYOF 928
Db 879 ANNLRQALQVRAGSHYAFSPMFEVLGQFVFEVRGSSRYNVLDLGGKQF 928
RESULT 6
AB990573
ID AB990573 standard; Protein; 928 AA.
XX
AC ABB90573;
XX
DT 29-JUL-2002 (first entry)
XX
DE Chlamydia pneumoniae cp0010 protein, SEQ ID NO:95.
XX
KW Chlamydial infection; antigen; immunogen; vaccine; diagnosis;
KW human respiratory disease; cardiovascular disease; atherosclerosis;
KW coronary artery disease; carotid artery stenosis; myocardial infarction;
KW cerebrovascular disease; aortic aneurysm; claudication; stroke;
XX strain CWL029.
OS Chlamydia pneumoniae.

XX	Key	Location/Qualifiers	
PH	Peptide	1..25	
FT	Protein	/label= Signal_peptide	
FT		26..328	
XX		/note= "Mature protein"	
XX		WO200202606-A2.	
XX		10-JAN-2002.	
XX		03-JUL-2001; 2001WO-IB01445.	
XX		03-JUL-2000; 2000GB-0016363.	
PR		11-JUL-2000; 2000GB-0017047.	
PR		21-JUL-2000; 2000GB-0017983.	
PR		07-AUG-2000; 2000GB-0019368.	
PR		18-AUG-2000; 2000GB-0020440.	
PR		14-SEP-2000; 2000GB-0022583.	
PR		10-NOV-2000; 2000GB-0027549.	
PR		22-DEC-2000; 2000GB-0031706.	
XX		(CHIR-) CHIRON SPA.	
XX		Ratti G, Grandi G;	
XX		WP1; 2002-154726/20.	
DR		N-PSDB; ABL91231.	
XX		Novel Chlamydia pneumoniae protein useful in the manufacture of a	
PT		medicament for treatment or prevention of infection due to Chlamydia,	
PT		preferably Chlamydia pneumoniae, and for diagnostic purposes -	
XX		Claim 1; Page 87-88; 364pp; English.	
XX		Sequences ABB90526-ABB90715 represent novel proteins from Chlamydia	
CC		pneumoniae (strain CWL029), and ABL91184-ABL91373 represent DNA encoding	
CC		them. The proteins are predicted to be immunogenic and may therefore be	
CC		useful in vaccine production and for diagnostic purposes. Chlamydia	
CC		pneumoniae is a common cause of respiratory disease in humans, and is	
CC		also involved in the development of cardiovascular diseases such as	
CC		atherosclerosis, coronary artery disease, carotid artery stenosis,	
CC		myocardial infarction, cerebrovascular disease, aortic aneurysm,	
CC		claudication and stroke. The proteins and nucleic acids of the invention	
CC		may be used in vaccines and pharmaceutical compositions for the	
CC		prevention or treatment of chlamydial infections, particularly Chlamydia	
CC		pneumoniae infections. The proteins may also be used in the detection of	
CC		Chlamydia pneumoniae, and the nucleic acids may be used in PCR, branched	
CC		DNA probe assay or blotting techniques for determining Chlamydia	
CC		pneumoniae gene expression. The present sequence represents a	
CC		specifically claimed Chlamydia pneumoniae protein of the invention.	
XX		Sequence 928 AA;	
SQ		Query Match 42.9%; Score 2048; DB 23; Length 928;	
		Best Local Similarity 46.5%; Pred. NO. 8.6e-126;	
		Matches 442; Conservative 144; Mismatches 320; Indels 44; Gaps 20;	
QY		1 MKSFPKPVF-STRAIF---PLSMATETVLDSASPDGKN-KNFSVRESQEDAGTTL 55	
DB		1 MKSQFSLVLSSTLACFTSCSTVFPAATENIGPDSFPGSINTGTYTPKNT--TTGIDYT 58	
QY		56 FKGNTVLENIPTGTATKSCFNNTKGLDITFTGNLSLLFQTVADAGTVGAAVNSVVDK 115	
DB		59 LTGDTITLQNL-GDSALFKGCFSDTTESLSFAGKGYLSFLNKS-SAEGAAL-SVTTDK 115	
QY		116 STTFTGFSLSFIASPGSIIT--KGAVSCSTGSLSTKVNLSLLFSKNFSTDNGGATA 173	
DB		116 NLSLTGFSLSLTFLAAPSSVITTPSGKAVKCG-GDLTFTDNNGTILFKODYCEBNGAIST 174	
QY		174 KTLSTLTGTMSTALFSENTSS---KKGALQTSDALTTITGNQCEVFSFNTSDSGAALFT 230	
DB		175 KNLSLKNSTGSGISFFGNKSSATGKKGKGAICATGTVDITNTTAPTFLFSNNIAEAGGAINS 234	
QY		231 EASVTISNNAKVSFIDNKVTCASSSTTCGMSGGALCAVKTSTDTKVTLTGQMLLFSNNT 290	
DB		235 TGNCTITGNTSLVFSNSVT---ATAG--NGGAL-----SGDADVTISNQSVTFSGNQ 283	
QY		291 STTAGGAIYVKKLELAS--GGLTLFSRNSVNGGTAPKGAIAIEDSGELSISADSGDIVF 348	
DB		284 AVANGGAIYAKKLTILASGGGVSPFLTIIVQTTAGNGGAILAAGECSLSAEAGDITF 343	
QY		349 LGNTVTSTTP-GTNRSSIDLTSKAKWTALRSAGRAIYFYDPIITGSSSTTVTVLKNET 407	
DB		344 NCGNAIVATTPTTKRNSIDIGSTAKITNLRAISGHSIFFYDPIITANTAADSTDTLNLAKA 403	
QY		408 PADSALOYTGNIIFTGEKLSETERADSKNLTSLKLPVTLTSGGTLSLKHGVTLQCAFTQ 467	
DB		404 DAGNSTDYSGSIVFSGEKLSEDEAKVADNLITSLQVTLTLAGNVLKRGVTLDTKGFTQ 463	
QY		468 QADSRLEMDVGTTLLEPA-DTSTINNLVINISSIDGAKKAKIETKATSKNLTLSGTITLLD 526	
DB		464 TAGSSVINDAGTTLKASTEEVTLTGLSIPVDSLGEKKVIAAASAASKNVALSGPILLD 523	
QY		527 PTGTFFYENHSLRNPOSYDILELKASGTVTSTAVTDPDPMGEKHFHYGQTWGPVW---- 582	
DB		524 NQGNAYENHDLGKTQDFSFVQLSALGTATTDDPAVPTVATETHYGYQGTWG-MTWVDDT 582	
QY		583 -GTGASTTATFNWTKGYIPNPERIGSLVPSNLWNAFIDISLHLYMETANEGLQGDRAF 641	
DB		583 ASTPKTKTATLAWNTGYLPNPERQGPLVPSNLWGSFSDIQAIOGVIERALSITLCSDRGF 642	
QY		642 WCAGLSNFFHKDSTKTRRGHLSGGYVIGNHLHTCSDKILSAAFCQLFGRDRDYFAKN 701	
DB		643 WAAGVANFLDKKKGKRYRHKSGGYAIGGAAQTCSENLSFQFQGLFGSKDKFLVAKN 702	
QY		702 QGTVYGGTLYYQHNETYIISLPCKLRPCSLSYVP---TEIPVLFSGNLSYTHTDNDLTKY 758	
DB		703 HTDYAGAFYIQH---ITECSGFIGCLLDKLPGSHKPLVLEGLQLAYSHVSNDLTKY 758	
QY		759 TTYPTVTKGSGNDSEALERFGGRAPICLDESALEQVYPMFKLQFVVAHQEGFKEQCTEAR 818	
DB		759 TAYPEVKGSGMGNFNMGLGASHSYPEVLHCFDTYAPYIKMLNTYIROQSFSEKTEGR 818	
QY		819 EFGSSRLNLALPIGIRFDKESCDQATYNLTLYTVDLVRSNPDCDCTTLRISGDSWKT 878	
DB		819 SFDDSNLNLSLPIGVKFEKSPDNDPSYDLTLYVVPDLIRNDPKCTTALVISGASWET 878	
QY		879 GTMLAQALVLRAGNHFCSNFSEAFSQSFELRGSSRNYNVDLGAQYQF 928	
DB		879 ANNLAQALQVRAGSHYAFSPMEVLQGFVFEVGRSSRIYNVDLGGKQFQF 928	
RESULT 7			
AA90237		ID AAY90237 standard; Protein; 928 AA.	
XX		AA90237;	
XX		29-AUG-2000 (first entry)	
DE		Chlamydia antigen CPN100635.	
KW		Chlamydia antigen; diagnosis; infection; community acquired pneumonia;	
KW		therapy; upper respiratory tract disease; bronchitis; sinusitis;	
KW		asthmatic bronchitis; adult-onset asthma; acute exacerbations of asthma.	
OS		Chlamydia pneumoniae.	
XX		Key Location/Qualifiers	
PH		1..43	
FT		/note= "signal peptide"	
FT		44..928	
FT		/note= "mature CPN100635"	
XX		WO200032794-A2.	

PI	Murkin AD, Omen RP;	
XX	WPI; 2000-224703/19.	
DR	N-PSDB; AA261509.	
XX	Novel antigens and corresponding DNA molecules that can be used to	
PT	prevent, treat and diagnose disease caused by Chlamydia infection in	
PT	mammals, especially humans -	
XX	Claim 19; Fig 15-E; 201pp; English.	
PS	AA9362-69 represent Chlamydia pneumoniae polypeptides. The	
XX	polypeptides are present in the bacterial membrane structure, in the	
CC	external vicinity of the membrane structure, in the inclusion membrane	
CC	structure, in the external vicinity of the inclusion membrane structure,	
CC	and in the cytoplasm of the infected cell. The polypeptides may be	
CC	used to prevent, treat and detect the presence of Chlamydia infection	
CC	and/or the presence of Chlamydia in a sample. The polypeptides may	
CC	also be used to induce an immune response in a mammal. The vaccine	
CC	vector comprising the polynucleotides is used to induce an immune	
CC	response in a mammal. Antibodies directed against the polypeptides	
CC	may also be used therapeutically to treat and/or prevent a Chlamydia	
CC	infection.	
XX	Sequence 918 AA;	
SQ	Query Match 42.0%; Score 2006; DB 21; Length 918;	
	Best Local Similarity 45.6%; Pred. No. 4.9e-123;	
	Matches 432; Conservative 150; Mismatches 316; Indels 50; Gaps 20;	
QY	1 MKSFPKVFSTAIPLSMI----ATETVLDSSAFDGN-KNGNSVRESQDA-GTTY 54	
Db	1 MRSPSLLISSLAPLMSVADAADLTGSRDSYNGDTSTETTPKAATSDAGTTY 60	
QY	55 LFGKGVNLENIPTGTATKSCFNNTKGDITFTGNGNSLTFQTVDAQTGAGAAVNSVVD 114	
Db	61 ILGDVSIQ-AKQVSLTSCFNNTAGNLTFLNGFSLHFDNIISSTVAGVVVNTAAS 119	
QY	115 KSTTFIGFSLSPASPSITGKGVSCSTGSLSTKVNLSLFSKXFTDNGGAIKAT 174	
Db	120 GITKFSGFSLMLAAR--TTGKAIKITDG-LVPEISGNLDNLNENASSENGGAINTK 175	
QY	175 TSLTGTMTALPSENTSSKKGAIQTSDALTITNGOGEVSPDNTSSDGAALFTASV 234	
Db	176 TSLTGTTRFVFLGNSSSQGGAIYASGVSISENAGILSPGNSSATSGGALSAGNL 235	
QY	235 TISNAKVSTIDNKVTGASSSTTGDMGGAIKAYK--TSTDTKVTLTGQNMQLFSNNST 292	
Db	236 VISNNQIFDGGCKAT-----TNGALDCNKAGANPDPIITLSGNSLHFLNNTAG 286	
QY	293 TAGGAIYVKKLELASG-GLTLFSPNSVNGTAPKGGAIATEDSGELSLSDSGDIVFLGN 351	
Db	287 NSGGAIYTKKLVLSRGGVLFSSNNKAANAT-PRGGAIALDLSGEISISADLGNIIPEGN 345	
QY	352 TVTSTT---PQTNSSIDLTSAKWTALRSAGRAIVFYDPIITGSGTITVDVLKVNETP 408	
Db	346 T-TSTTGPASVTRNADLNAKFLNLRATRGKVIIFYDIT---SSGATDKLSLNKAD 401	
QY	409 ADSALQVTGNIIFTGEKLSATEAADSKNLTSKLPQVTLSGGTLSLKHGVTLOQAFQQ 469	
Db	402 AGSGNTVEGYIVFSGEKLSEELKPNLKSFTQAVELAAGALVLDKGVTVVANTITQV 461	
QY	469 ADSLEMDVGTLE-PADTSTINNLVINISIDAKAKIETKATSKNLTLSGTITLDP 527	
Db	462 EGSKVMDGGTTFEASBEGVTGLNIAINIDLDGTNKAIIKATASKDVALSGPIMLVDA 521	
QY	528 TGTVEHNSLNPQSDILELKAQSTVTSTVTPDPIGMEFHYGYQGTWGPYV---GT 584	
Db	522 QGNYEHEHLSQQQVFFIILSAQGTMTTIDPITPILNTNHHYGYQGNW-IVVDDAT 580	
QY	585 GASTTATENWTKGTGYPNPERIGSLVPSNLNAFIDISLHYLMETANEGLQGDRAFWCA 644	
Db	581 AKTKNATLTWTKYTKPNRQGPPLVPNSLWSGFDVVERISQSLMDRSFSSLSSTNLWVS 640	

QY	645 GLSNFFHKDSTKTRGRHRLSGGYVIGGNLHLCSDKILSAAPCOLFGRRDRDYFVAKNGT 704	
Db	641 GIADFLHEDQKGNQRSYRSHSAGYALGGGFTASENFNFAPCOLFGYDKDHLVAKNHTH 700	
QY	705 VYGGLTYQH---NETYISLPCKLKRLPCSLSVYPTPIVLFSGNLSYTHDNLTKYTTY 761	
Db	701 VYAGAMSYRHLGESKTLAKI-----LSGNSDSLFPVFNARPAYGHTDNNMTKYTG 752	
QY	762 PTVKSGWNSDFALPFGGRAPICLD-ESALFEQYMPFMKLOFVYAHQSGFKQCTEAREF 820	
Db	753 SPVKSGWNGDAFGIECGGAIFVVASGRRSWVDTHPFLNLEMIYAHQNDFKENGTEGRSF 812	
QY	821 GSSRLVNLALPIGIRFDKESQCDQATNLIGYTVDLVRSNPDCTTTLRISGDSMKTFT 880	
Db	813 QSEDLFNLAVPVGIRKFXSD--KSTYDLSTAYVDPVIRNDPCGCTTLLMVGSDSWSTCGT 870	
QY	881 NLARQALVLRAGNHCFNNSFEAFSQSFELRGSSRNYNVDLGAKYQF 928	
Db	871 SLSRQALLVRAGNHAFASNPFEVFSQFEVLRGSSRSYDAIDLGGRFGF 918	
XX	RESULT 9	
ID	AAW88422 standard; Protein; 918 AA.	
AC	AAW88422;	
XX	26-APR-1999 (first entry)	
XX	Chlamydia pneumoniae surface exposed protein Omp9.	
XX	Omp9; outer membrane protein 9; surface exposed protein; antigen;	
KW	infection; diagnosis; vaccine; atherosclerosis; asthma.	
XX	Chlamydia pneumoniae.	
OS	WO9858953-A2.	
PN	30-DEC-1998.	
PD	19-JUN-1998; 98WO-DK00266.	
XX	23-JUN-1997; 97DK-0000744.	
PF	(BIRK/) BIRKELUND S.	
XX	(CHRI/) CHRISTIANSEN G.	
PA	Birkelund S, Christiansen G, Knudsen K, Madsen A;	
XX	Mygind P;	
XX	WPI; 1999-105610/09.	
DR	N-PSDB; AA06821.	
XX	Species-specific test for identifying mammals infected with	
PT	Chlamydia pneumoniae - comprises detecting antibodies specific for	
PT	outer membrane proteins of C. pneumoniae or nucleic acids encoding	
PT	these proteins	
XX	Claim 7; Page 56-58; 115pp; English.	
XX	This polypeptide comprises the novel 96.7 kDa surface exposed	
CC	protein Omp9 of the human respiratory pathogen Chlamydia	
CC	pneumoniae. Its amino acid sequence was deduced from DNA (see	
CC	AA06821) isolated from a C. pneumoniae expression library. The	
CC	invention provides 12 novel surface exposed proteins, Omp4-Omp15	
CC	(see AA068417-28), and nucleic acid sequences encoding them (see	
CC	AA06816-27). A new species specific test is claimed that is used	
CC	to identify mammals (including humans) infected with Chlamydia	
CC	pneumoniae. The test comprises detecting antibodies specific for	
CC	Omp4-Omp15 or detecting nucleic acid fragments encoding these outer	
CC	membrane proteins, especially by PCR. The proteins are also used	
CC	in the diagnosis of C. pneumoniae infection in mammals. The	

Db 116 NLSLTGFFSLTFLAAPSIVITPSGKGVKCG-GDLTFDNNGTILFKQDYCEENGGAIST 174
Qy 174 KILSLTGTTMSALFSENTSS---KKGAIQTSALITGNOGEVFSFSDNTSSDGAIF 230
Db 175 KNLSEKNSGTSISFEGNKSATGKGGAI-CATGTVDTTNTAPTLESNNIAAAGGAINS 234
Qy 231 EASVTISNNAKVSFIDNKVTGASSSTTGDMSGGAI-CAYKTSTDTKVTLLTGNNQMLFNSNT 290
Db 235 TGNCTTGTNTSLVSENSVT---ATAG--NGGAL-----SGDADVTISGNSQSVTFSGNQ 283
Qy 291 STTAGGAIYVKKLELAS--GGLTLFSRNSVNGTAPKGGAIATEDSGELSLNDSGDIVF 348
Db 284 AVANGGAIYAKKLTLAGGGGGISFNNIVQGTAGNGGAIISLAAGECSLSAEGDIIF 343
Qy 349 LGNTVSTTP--GNRRSIDLGTAKMTALRSAGRAIYFYDPTTGGSTTIVTDLVKVNET 407
Db 344 NGNAIVATTPTQTKRNSIDIGSTAKITNLRAISGHSIFFYDPTITANTAADSTDTLNKA 403
Qy 408 PADSALOYTGNIIFTGKLESETBAADSKNLTSLKLPVLSGGTSLKHGVTLOTAFTQ 467
Db 404 DAGNSTDYSGSIVFSGEKLSEDAKVADNLTSLKQPVTLTAGNLVLRGVTLTKGFTQ 463
Qy 468 QADSRLEMDVGTILEPA-DTSTINNLVINISSIDGAKAKIETKATSKNLTLSGTTILLD 526
Db 464 TAGSSVIMDAGTTKASTEEVTLTGLSIPVDSLGEKKVIVIAASAASKNVALSPILLLD 523
Qy 527 PTGTFYENHSLRNQSDILELKAAGTGTATVTPDPIMGEKPHYGYQGTWGPVW--- 582
Db 524 NQWAVENHDLGKTQDFSVQLSALGTATTVDPAVTPATPHYGYQGTW--MTWDDT 582
Qy 583 -GTGAATTATFNWTKTYIIPNPERIGSLVPSNLSNAPFIDISLHLYMETANEGLQGDRAF 641
Db 583 ASTEKTATLATWNTGYLPNPGRQGPLVPSNLMSGSPSDIQAIGQVIERSAITLCSDRGF 642
Qy 642 WCAGLSNPFHKDSTKTRGRHLSGGYVIGNLHTCSDKILSAFCOLFGRDRDYFAKN 701
Db 643 WAAGVANFLDKDKKGRKRYRHKSGGYAIGGAAGTCSENLIISFAFCQLFGSDKDFLVAKN 702
Qy 702 QGTVYGTLVYQHNETVSLPCKLRPCSLSVYP---TEIPVLFSGNLSYTHTDNDLTKY 758
Db 703 HTDYAGAFYIQH-----ITECSGFIGCLLDKLPQSWHKLPLVLEGQLAYSHVSNLTKY 758
Qy 759 TTYPTVKGWGNDSFALEFGGRAPICLDESALFPQYMPFMKLFQYAHQEGFKQGTAR 818
Db 759 TAYPEVKGWGNNAFNMMLGASSHSYPEYLHCFDTYAPYIKNLTYIRQDSFSEKGTGR 818
Qy 819 BFGSSRLVNLALPTGIRFKESDCODATYNLTGLVTVDLVRSNPDCTTLIRISGDSWTF 878
Db 819 SFDDSNLFLSLPIGVKFEKFCDCNDFSYDTLTSVYDPLIRKNDPKCITVALVISASWET 878
Qy 879 GTNLARQALVLRAGNHFCFNSNFAPFSQFSELRGS 914
Db 879 ANNARQALQVRAGSHYAFSPMFEVLGQFVEVRGS 914

RESULT 11

AA90238
ID AA90238 standard; Protein; 885 AA.
XX AA90238;
AC AA90238;
XX
XX 29-AUG-2000 (first entry)
XX Mature Chlamydia antigen CPN100635.
DE Chlamydia antigen; diagnosis; infection; community acquired pneumonia;
KW therapy; upper respiratory tract disease; bronchitis; sinusitis;
KW asthmatic bronchitis; adult-onset asthma; acute exacerbations of asthma.
OS Chlamydia pneumoniae.
XX
XX WO200032794-A2.

XX PD 08-JUN-2000.
XX
XX 01-DEC-1999; 9990-CA01147.
XX
XX 01-DEC-1998; 98US-0110339.
PR 01-DEC-1998; 98US-0110340.
PR 01-DEC-1998; 98US-0110427.
PR 01-DEC-1998; 98US-0110428.
PR 01-DEC-1998; 98US-0110438.
XX (CONN-) CONNAUGHT LAB LTD.
XX
XX Murdin AD, Omen RP, Wang J;
PI
XX
XX WPI; 2000-412339/35.
DR N-PSDB; AAA30849, AAA30850.
XX
XX Nucleic acids encoding polypeptide antigens from Chlamydia useful for preventing, diagnosing and treating diseases such as community acquired pneumonia, bronchitis, sinusitis and asthmatic bronchitis, adult-onset asthma -
XX
XX Claim 16; Fig 3; 174pp; English.
XX
XX This sequence is a Chlamydia antigen of the invention, designated CPN100635. The nucleic acids (and their complementary sequences) may be used as diagnostic agents for detecting the presence of nucleic acids encoding Chlamydia antigens in samples according to standard methods, and therefore, for diagnosing Chlamydia infections. For example, they may be used as primers and probes for diagnostic polymerase chain reaction (PCR) assays. Antisense sequences may be used to down regulate expression of the proteins and may be used to treat infections. The nucleic acids may also be used to produce the protein antigens they encode according to standard recombinant DNA methodologies. The proteins may then be used as antigens for the production of antibodies (i.e. as vaccines) for preventing infection by Chlamydia. The antibodies may also be used as diagnostic reagents for detecting infections. Chlamydia is a pathogen implicated in the development of (for example) community acquired pneumonia, upper respiratory tract disease (especially bronchitis and sinusitis), asthmatic bronchitis, adult-onset asthma and acute exacerbations of asthma in adults.
XX
XX Sequence 885 AA;

Query Match 41.6%; Score 1987.5; DB 21; Length 885;
Best Local Similarity 46.9%; Pred. No. 7.6e-122;
Matches 420; Conservative 136; Mismatches 302; Indels 37; Gaps 16;
Qy 51 GTTYLPKGNVTLENIPTGTATKSCFNNTKGDLTFTGNGNSLLFQTVDA GTVAGAAVNS 110
Db 11 GIDYTLTGDTLQNL-GDSALTKGCFSDTTESLPAGKGYSLFLNKS-SABGAAL-S 67
Qy 111 SVVDKSTTFIGFSSLSFIASPGSSITT--GKGAVSCSTGSLSTKNVLSLFKNFSTNDG 168
Db 68 VTTDKNLSLTCGFSLSLTLAAPSIVITPSGKGVKCG-GDLTFDNNGTILFKQDYCEENG 126
Qy 169 GAITAKTSLTGTTMSALFSENTSS---KKGAIQTSALITGNOGEVFSFSDNTSSDGS 225
Db 127 GAISTKNLSKNSTGTSISFEGNKSATGKGGAI-CATGTVDTTNTAPTLESNNIAAAG 186
Qy 226 AAIETASVTISNNAKVSFIDNKVTGASSSTTGDMSGGAI-CAYKTSTDTKVTLLTGNNQML 285
Db 187 GAINSTGNTCTTGTNTSLVSENSVT---ATAG--NGGAL-----SGDADVTISGNSQSV 235
Qy 286 FSNNTSTTAGGAIYVKKLELAS--GGLTLFSRNSVNGTAPKGGAIATEDSGELSLNDS 343
Db 236 FSGNQAVANGGAIYAKKLTLAGGGGGISFNNIVQGTAGNGGAIISLAAGECSLSFSEA 295
Qy 344 GDIVFLGNTVSTTP--GNRRSIDLGTAKMTALRSAGRAIYFYDPTTGGSTTIVTDL 402
Db 296 GDHYLNGNAIVATTPTQTKRNSIDIGSTGKDHLELRALISGHSIFFYDPTITANTAADSTDL 355

645 EGISNFFHKDSTKINKGFRHISAGYVVGATTTLASDNLITAAFCQLFGKDRDHFINKRA 704
704 TVYGGTLYYQHNETYISLPCKLRPCSLSYVP---TEIPVPSGNLSYTHTDNDLTKYTT 760
705 SAYAASLHLQHIAIATLSS-----PSLLRYLPQSESEQPVLPDAQISYIYKNTMYTYTQ 758
761 YPTVKGSGWGNDSFALEFGGRAP-ICLDESALFEQYMPFMKLOFVYAHQSGFKEQGTG-AR 818
759 APKGSSWYNDGCALELASSLPHTALSHEGLFHAYFPFIKVEASYIHQDSFKERNITLVR 818
819 EFGSSRLVNLALPIGIRPKESQCDATNLTGTYVDLVRSPDCQTTTLRISGSKWTF 878
819 SFDSSGLINVSVPFIGHTFRFSNERASYEATVIYVADYVRKNPDCTTALLINNTSWKTT 878
879 GTNLARQALVLRAGNHFCFNSNEAFSQFSFELRGSSRNYNVDLGAQYQF 928
879 GTNLSRQAGIGRAGICYAFSPNLEVTNLSMEIRGSSRSYNADLGKQFQF 928

RESULT 13

ID ABB90542 standard; Protein; 928 AA.

XX AC ABB90542;
XX DT 29-JUL-2002 (first entry)

Chlamydia pneumoniae cp6731 protein, SEQ ID NO:33.

XX Chlamydial infection; antigen; immunogen; vaccine; diagnosis;
KW human respiratory disease; cardiovascular disease; atherosclerosis;
KW coronary artery disease; carotid artery stenosis; myocardial infarction;
KW cerebrovascular disease; aortic aneurysm; claudication; stroke;
XX strain CML029.

XX Chlamydia pneumoniae.

XX Key Location/Qualifiers
FH Peptide 1..26
FT /label= Signal_peptide
FT Protein 27..928
FT /note= "Mature protein"

WO200202606-A2.

10-JAN-2002.

03-JUL-2001; 2001WO-IB01445.

03-JUL-2000; 2000GB-0016363.
11-JUL-2000; 2000GB-0017047.
21-JUL-2000; 2000GB-0017983.
07-AUG-2000; 2000GB-0019368.
18-AUG-2000; 2000GB-0020440.
14-SEP-2000; 2000GB-0022583.
10-NOV-2000; 2000GB-0027549.
22-DEC-2000; 2000GB-0031706.

(CHIR-) CHIRON SPA.

Ratti G, Grandi G;

WPI; 2002-154726/20.

N-PSDB; ABL91200.

XX Novel Chlamydia pneumoniae protein useful in the manufacture of a
PT medicament for treatment or prevention of infection due to Chlamydia,
PT preferably Chlamydia pneumoniae, and for diagnostic purposes

XX Claim 1; Page 57; 364pp; English.

XX Sequences ABB90526-ABB90715 represent novel proteins from Chlamydia
XX pneumoniae (strain CML029), and ABL91184-ABL91373 represent DNA encoding

CC then. The proteins are predicted to be immunogenic and may therefore be
CC useful in vaccine production and for diagnostic purposes. Chlamydia
CC pneumoniae is a common cause of respiratory disease in humans, and is
CC also involved in the development of cardiovascular diseases such as
CC atherosclerosis, coronary artery disease, carotid artery stenosis,
CC myocardial infarction, cerebrovascular disease, aortic aneurysm,
CC claudication and stroke. The proteins and nucleic acids of the invention
CC may be used in vaccines and pharmaceutical compositions for the
CC prevention or treatment of chlamydial infections, particularly Chlamydia
CC pneumoniae infections. The proteins may also be used in the detection of
CC Chlamydia pneumoniae, and the nucleic acids may be used in PCR, branched
CC DNA probe assay or blotting techniques for determining Chlamydia
CC pneumoniae gene expression. The present sequence represents a
CC specifically claimed Chlamydia pneumoniae protein of the invention.

XX Sequence 928 AA;

Query Match 41.5%; Score 1982; DB 23; Length 928;

Best Local Similarity 45.2%; Pred. No. 1.9e-121;

Matches 429; Conservative 150; Mismatches 327; Indels 44; Gaps 19;

QY 1 MKSSPPKVFVSTFAIPPLSM-----IATETVLDSSASPDGNKNGNFVSRESQEDA-GTT 53

Db 1 MKSSLHWFLISSSLALPLSLNFSFAFAAVVEINLGPNTSPSG--PGTYTPPAQTNNADGTI 58

QY 54 YLPKGNVLENIPTGCTAITKSCENNTKGDLTFTGNGNSLLFTQVDVAGVAAVNSSVY 113

Db 59 YNLTDGVSIITN-AGSPALTATSCFETTNLNSFGQHGQYFLQNDIAG--ANCTFNTAA 115

QY 114 DKSTTFEIGFSSLSFIASPGSSITTKGAVSCSTGSLSTKNVLSLFKSNFSTDNGGAI 173

Db 116 NKLLSFGSYLSLIQT--TNATTGGAIK-STGACSIQSNYSYCPQNFSDNGGALQG 172

QY 174 KILSLGTMTMSALFSENSTSKKGAIQTSALTITGNGEVSPSNTSSDGAITEAS 233

Db 173 SSISLS-LNPNTLFAKNKATOKGALYSTGGITINNTLSASPSSENTAANGGAIYTEAS 231

QY 234 VTISNNKVSFIDNKVYTGASSSTTGDMSGCAICAYKTTSTDTKV-TLTGNCMLFSSNNTST 292

Db 232 SFISSNKAFSFINNSVTATSA-----TGGAIVCSSTSPKPVLTISDNGELNFIGNTAI 285

QY 293 TAGGAIYVKKLELASGGLTLFSRNVNGGTAPKGGAIATBDGSELSDSGSDIVFLGNT 352

Db 286 TSGGAIYTDNLVLSGGPTLTKNSALDTAALPGGAIADSGSLSLALGGDITEGNT 345

QY 353 V---TSTTPGTNRSSIDLG-TSAKMTALRSAAGRAIYFDYDPTTGTSTTTVDLKNETP 408

Db 346 VVRGASSQTTTRNSINIGNTNAKIVQLRASQGNITFYFDPTTTSITAAUSDALNNGPD 405

QY 409 ADSALQVTGNIIFTGKELSETEADSKNLTKLQPVTLGGTSLSKHGVTLQTAFTQQ 468

Db 406 LAGNPAYQGIIVFSGEKLSEAEAEADNLKSTTQQPTLAGGQLSKSGVTLVAKSFSQS 465

QY 469 ADSREMDVGTTLPEADTSTINNLVINISSIDGAKAKIETKATSKNLTLSGITLLDPT 528

Db 466 PGSTLLMDAGTTLTETADGITINNVLNVDSLSLTKETKATLQATQASQVTLSSLSLVDPS 525

QY 529 GTFYENHSLRNPQSYDILELKAS--GTVTSTAVTPDPIMGEKPHYGYQGTWGPVWGTGA 586

Db 526 GNVYEDVSWNPNQVFSCLTTLTADDPANIHTDLAADPLEKNPIHWGQGNWA-LSWQEDT 584

QY 587 ST---TATENWTKGYIPNPERIGSLVPNSLNWAFIDISSLHYLMETANEGLCQDRAFWC 643

Db 585 ATKSKAATLWTKTGYNPNPERRGTILVANILWGSFVDRSIQQLVAIKVRSQSETRGWC 644

QY 644 AGLSNFFHKDSTKTRRGFRHLGSGYVIGGNLHTCSDKILSAAPCOLFGRDRDYFAVKNQG 703

Db 645 EGISNFFHKDSTKINKGFRHISAGYVVGATTTLASDNLITAAFCQLFGKDRDHFINKRA 704

QY 704 TVYGGTLYYQHNETYISLPCKLRPCSLSYVP---TEIPVPSGNLSYTHTDNDLTKYTT 760

Db 705 SAYAASLHLQHIAIATLSS-----PSLLRYLPQSESEQPVLPDAQISYIYKNTMYTYTQ 758

QY 761 YPTVKGWNSDFALEPFGGRAP-ICLDBSALFEQYMPFMKLOFYVAHQEGKEQGT-AR 818
 Db 759 APKGSSWYNDGCALASSLPHLTAHSHEGLFHFAYFPFIKVEASYIHQDSFKERNITLVR 818
 QY 819 EFGSRLVNLALPIGIRPKESDCODATYNLTGTVDLVRGNPDCTTTLRISGDSWKT 878
 Db 819 SFDSDGLINVSPIGIFERFSRNERASYEAVIYVADYVRKNPDCTTALLINNTSWKT 878
 QY 879 GTNLARQALVLRAGNHFCSNFEPFQSFRLGSSRNYNVDLGAQYOF 928
 Db 879 GTNLARQALVLRAGNHFCSNFEPFQSFRLGSSRNYNVDLGAQYOF 928

RESULT 14

AAV90239
 ID AAY90239 standard; Protein; 928 AA.

AC AAY90239;

DT 29-AUG-2000 (first entry)

XX Chlamydia antigen CPN100638.

XX Chlamydia antigen; diagnosis; infection; community acquired pneumonia;
 KW therapy; upper respiratory tract disease; bronchitis; sinusitis;
 KW asthmatic bronchitis; adult-onset asthma; acute exacerbations of asthma.

XX Chlamydia pneumoniae.

XX W0200032794-A2.

PN 08-JUN-2000.

XX 01-DEC-1999; 99WO-CA01147.

PF 01-DEC-1998; 98US-0110339.

PR 01-DEC-1998; 98US-0110340.

PR 01-DEC-1998; 98US-0110427.

PR 01-DEC-1998; 98US-0110428.

PR 01-DEC-1998; 98US-0110438.

XX (CONN-) CONNAUGHT LAB LTD.

PA Murdin AD, Oomen RP, Wang J;

PI WPI; 2000-412339/35.

DR N-PSDB; AAA30851, AAA30852.

XX Nucleic acids encoding polypeptide antigens from Chlamydia useful for

PT preventing, diagnosing and treating diseases such as community acquired
 PT pneumonia, bronchitis, sinusitis and asthmatic bronchitis, adult-onset
 PT asthma -

XX Claim 16; Fig 5; 174pp; English.

XX This sequence is a Chlamydia antigen of the invention, designated
 CC CPN100638. The nucleic acids (and their complementary sequences) may be
 CC used as diagnostic agents for detecting the presence of nucleic acids
 CC encoding Chlamydia antigens in samples according to standard methods,
 CC and therefore, for diagnosing Chlamydia infections. For example, they may
 CC be used as primers and probes for diagnostic polymerase chain reaction
 CC (PCR) assays. Antisense sequences may be used to down regulate
 CC expression of the proteins and may be used to treat infections. The
 CC nucleic acids may also be used to produce the protein antigens they
 CC encode according to standard recombinant DNA methodologies. The
 CC proteins may then be used as antigens for the production of antibodies
 CC (i.e. as vaccines) for preventing infection by Chlamydia. The
 CC antibodies may also be used as diagnostic reagents for detecting
 CC infections. Chlamydia is a pathogen implicated in the development of
 CC (for example) community acquired pneumonia, upper respiratory tract
 CC disease (especially bronchitis and sinusitis, asthmatic bronchitis,
 CC adult-onset asthma and acute exacerbations of asthma in adults.

SQ Sequence 928 AA;

Query Match 41.4%; Score 1976; DB 21; Length 928;

Best Local Similarity 45.1%; Pred. No. 4.6e-121;

Matches 428; Conservative 150; Mismatches 328; Indels 44; Gaps 19;

QY 1 MKSSPPKVFSTFAIFPLSM-----IATETVLSSASFQGNKGNFVSRESQEDA-GTT 53

Db 1 MKSSLHFLISSLLALPLSLNFAFAAVVEINLFTNSFG--PGTYPPPAQTWADGTI 58

QY 54 YLFKGNVTLENIPTGTATITKSCFNNTKGDLTFTNGNSLLFPQTVDAGTVAGAAVNSV 113

Db 59 YNLTDGVSI TN-AGSPTALTATSCFKETITGNLSFQGHYQFLQNLIDAG--ANCTFTTAA 115

QY 114 DKSTTFIGFSSLSFIASPGSSITTCGKGVSCSTGSLTKNVLSLFSKNFSTDNCGAITA 173

Db 116 NKLSFGSGFSYLSLIQT--TNATTGTGAIK-STGACISQSNYSCTYFGQNFNDNGALQG 172

QY 174 KTLSLTGTMTSALFSENTSSKKGGAIOISDALITIGNQGEVFSFNTSSDGAALFTEAS 233

Db 173 SSISLS-LNPNLTFKKNKATQKGAISTGGITINNTLSASFSENTAANGGAIVTEAS 231

QY 234 VTISNAKVSTFDNKTGVTGASSTTGDMSGGAICAYKTSTDTKV-TLTGNQMLLFSNIST 292

Db 232 SPISNNKAISFINNSVTATSA-----TGGAIVCSSTSAKPVLTLSNDGELNFIQNTAI 285

QY 293 TAGGAIYVKKLELAGSLTLPSENUNGCTAPKGAIALEDSELSADSGDIVFLGNT 352

Db 286 TSGGAIYTDNLVLSGGPTLFRNNGYDTAAIPGGGAIAADSGLSLALGGDITPEGNT 345

QY 353 V---TSTPTGTRNSSIDLG-TSAKMTALSAAGRAIFYDYDPTTGGSTTTVDLVKNETP 408

Db 346 VYKGASSQTTFNSINIGTNAKIVQLRASQNTIIFYDYDPTTSTTAALSADALNNGPD 405

QY 409 ADSALQYTCNIIPTGEKLTSETAADSKNLTSLKLOVTLGGTSLKHGVTLOTQAPTQQ 468

Db 406 LAGNPAYQGTIVFSGEKLSEAAEADNLKSTIQPLTLAGGQLSKSGVTLVAKFSOS 465

QY 469 ADSRLQMDVGTTLLEPADTSTINNLVINISIDGAKAKIETKATSKNLTSGTITLLDPT 528

Db 466 PGSTLLMDAGTTLETADGITINNLVNDLSLKEKTKGTLKATQASQTVTLGSLSLVDPS 525

QY 529 GTFYENHSLRNPOSYDILELKAS--GTVTSTAVTPDIPMGEKPHYGYQGTGWGIVMGTC 586

Db 526 GNVEDVSWNNPVQVFSCLTTLADDPANIHITLADPLEKNPIHWGYQGNWA-LSHQEDT 584

QY 587 ST---TATFNWTKTYIPNPERIGSLVPNSLNAFIDISLHYLMETANEGLDQDRAFWC 643

Db 585 ATKSKAATLTWTKTYNPNPERRGTIVANTLWGSFVDVRSIQQLVATKVRQSQETRGIMC 644

QY 644 AGLSNFPHKDSKTRGRFHLGSGYVIGGNLHCTCDKILSAAFCOLPGRDRDYFVAKNOG 703

Db 645 EGISNPFHKDSTKINKGFRHISAGYVVGATTTLASDNLITAAFCQLFGKDRDHFINKRA 704

QY 704 TVYGGTLYYQHNETYISLPCKLPFCSLYVP---TEIPVLFSGNLSYTHTDNDLKTYYT 760

Db 705 SAYAASHLQHLATLSS-----PSLLRYLPQSESEQPVLFDAQISVIYSKNTMKTYYTQ 758

QY 761 YPTVKGWNSDFALEPFGGRAP-ICLDBSALFEQYMPFMKLOFYVAHQEGKEQGT-AR 818

Db 759 APKGSSWYNDGCALASSLPHLTAHSHEGLFHFAYFPFIKVEASYIHQDSFKERNITLVR 818

QY 819 EFGSRLVNLALPIGIRPKESDCODATYNLTGTVDLVRGNPDCTTTLRISGDSWKT 878

Db 819 SFDSDGLINVSPIGIFERFSRNERASYEAVIYVADYVRKNPDCTTALLINNTSWKT 878

QY 879 GTNLARQALVLRAGNHFCSNFEPFQSFRLGSSRNYNVDLGAQYOF 928

Db 879 GTNLARQALVLRAGNHFCSNFEPFQSFRLGSSRNYNVDLGAQYOF 928

RESULT 15
 AAY99842

ID AAY99842 standard; Protein: 936 AA.
 AC AAY99842;
 XX
 XX
 XX 15-SEP-2000 (first entry)
 DE Chlamydia pneumoniae 98 kDa outer membrane protein CPN100640.
 DE
 DE Chlamydia, 98 kDa outer membrane protein; antigen; immunogen; infection;
 KW vaccine; antibacterial; community acquired pneumonia; bronchitis;
 KW sinusitis; acute respiratory disease; upper respiratory tract disease;
 KW asthma; atherosclerosis.
 XX
 XX
 XX Chlamydia pneumoniae.
 OS
 XX
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..11
 FT /label= Signal_peptide
 FT Protein 12..936
 FT /label= 98_kDa_membrane_protein
 XX
 PN W0200032784-A1.
 XX
 XX 08-JUN-2000.
 PD
 XX
 XX 01-DEC-1999; 99WO-CA01148.
 PF
 XX
 XX 01-DEC-1998; 98US-0110439.
 PR
 XX 03-MAY-1999; 99US-0132272.
 PR
 XX
 XX (CONN-) CONNAUGHT LAB LTD.
 PA
 XX
 XX
 XX Murdin AD, Oomen RP, Wang J, Dunn P;
 PI
 XX
 DR WPI; 2000-412330/35.
 DR N-PSDB; AAA48838, AAA48839.
 XX
 XX New polynucleotide encoding the Chlamydia 98 kilobalton outer membrane
 PT protein, useful for preventing or treating Chlamydia infection -
 XX
 XX Claim 16; Fig 1; 98pp; English.
 PS
 XX
 CC The present sequence is CPN100640, the 98 kDa outer membrane protein
 CC from Chlamydia pneumoniae. Chlamydia pneumoniae is a common cause of
 CC community acquired pneumonia and upper respiratory tract symptoms and
 CC diseases, including bronchitis and sinusitis. It also has an association
 CC with atherosclerosis and asthma. The 98 kDa outer membrane protein is a
 CC C. pneumoniae-specific antigen which can confer immune protection against
 CC Chlamydia infection. The nucleotide sequence encoding the protein or the
 CC protein itself may be administered as a vaccine to prevent or treat
 CC infection and they may also be used to diagnose infection. The gene
 CC encoding CPN100640 was amplified from Chlamydia pneumoniae genomic DNA
 CC by PCR.
 XX
 XX
 SQ Sequence 936 AA;
 Query Match 41.2%; Score 1965; DB 21; Length 936;
 Best Local Similarity 43.4%; Pred. No. 2.5e-120;
 Matches 411; Conservative 179; Mismatches 327; Indels 30; Gaps 16;
 QY 1 MKSSFPKVFVSTFAIF-PLSMIATETVLDS--ASFDGKNGNFSVRESOE-DAGTYTLFK 57
 DB 1 MKSSVSWLFFSIPPLFSSLSIVAAEVLDSNNYDGSNGTTFVFTDAAAGTYSLL 60
 QY 58 GNVTLNIPGTGTAITKSCFNNTKGDLTFTGNSNLLFTQVDAGTVAGAVNSVVDKST 117
 DB 61 SDVSFONAGALGIPLASGCGFLEAGDGLTPQGNQHALKFAFINAGSAGTAVSTAADKRL 120
 QY 118 TFIGSSLSFIASPGSSIT-TGKGVSCSTGSLSLTKNVSLFLSKNFTDNGCAITAKTL 176
 DB 121 LFNDFSRSLIISCPILLSPTGQCALK-SVGNLSLTGNQIIIFTQNFSSDNGGVINTKNF 179
 QY 177 SLTGTMTSALFSENT--SSKKGAIQTSDALITITGNQGEVFSFSDNTSDSGAAIFTEASV 234

Db 180 LLSGTSQFASFRNQAFQTKQGGVYATGTITTIENSPIVFSQNLAKSGGALSTDMC 239
 QY 235 TISNNAKVSFIDNKVTGASSTTGDMSGGAI CAYKTSTDTKVTLTGNQMLFNNSTTFA 294
 Db 240 SITDNFQVIFDCNSAWEAAQA-----QGGAI CC--TTTDTKVTLTGNKLSFTNTALT 292
 QY 295 GGAIVYKULEASGGTLTFSRNSVNGGTAPK--GGAIAIEDSGELSLSDSGDIVPLGNT 352
 Db 293 GGAISGLKVSISAGGPTLFQSN-IGSSAGQGGGAINIASAGELALSATSGDITFNNQ 351
 QY 353 VTSPTGTRSSITDLGTSAKMTALRSAGRAIYFYDPTTGGSTTVDLVKNVETPADSA 412
 Db 352 VTNGSTST-RNAINIIDTAKVTSIRATQCSYFYDPTINPGTAOSTDTLNLDANSE 410
 QY 413 LQVTGNIIFTGKLSETEAAADSKNLTSLKLPVTLGGTSLSKHGVTLTQTQAFQADSR 472
 Db 411 IEYGAIVFSGEKLSPTKAIANVTSTIRQPAVLARGDLVLRDGVTVTFKDLTQSPGR 470
 QY 473 LEMDVOTTLBPADTS-TINLVINISSIDGAKAKIETKATSKNLTSLGTTLLDPTGTF 531
 Db 471 ILMDGGTTLISAKEANLSLGLAVNLSLSDGTNKAALKTEAADKNISLSGTIALIDTEGSF 530
 QY 532 YENHSLRNQSYDILELK--ASGTVTSTAVTDPIMGEKPHYGYOQTWGPVWGTGAST 588
 Db 531 YENHLSKASTYPLELLETTAGANGTITLQEPETHYGYQGNW-QLSWANATSS 589
 QY 589 -TATFNWTKTYIPNPERIGSLVPSNLWNAFIDISLHYLMETANEGLOQDRAFWCAGLS 647
 Db 590 KIGSINWTRTYIPSPERKSNLPLNSLWGNFIDIRSIINQLIETKSSGEPERELMLSGIA 649
 QY 648 NPFHKDSTKTRRGFRHLSCGYVIGGNLHTCSDKILSNAFCOLFGRDRDYFVAKNQGVY 707
 Db 650 NFFYRDSMPTRHGRHISGGYALGITTATPAEDQLTFACQLFARDNRNHTKNGHDYTG 709
 QY 708 GTLYYQHNETYISLPCKL-----RPCSLSVVTEIPVLPFSGNLSYTHTDNDLKYTTY 761
 Db 710 ASLYFPHTEGLFDIANFLWKGKATPAWVLSIQIPLSFDKAFSYLHTDNHMKTYTDN 769
 QY 762 PTYKSGMGNDSFALFQGRAPICLDESFALFEQWMPFKLOFVYAHQEGFKEQGTAREFG 821
 Db 770 SIIRGWRNDAFCADLGLASLPFVISVPLYLKEVEPFVKQVYIYAHQODFYERHAEGRAF 829
 QY 822 SSRILVNALDIGIRFDKESDCQDATYNTLTGYVDLVRSNPDCTTTLRIISGDSWKTGTN 881
 Db 830 KSELINVEIPIGVTTFERDSKSEKGYDGLTILMYLDAYRNPCKOTSLIADANMMAYGTN 889
 QY 882 LARQALVLRAGNHPFCFNSFPAFQSFELRGSRNRYNVDLGAQYF 928
 Db 890 LARQGSVRAANHPQVNPHEIFCQAFEFVRSRNRNTNLGSKFCF 936

Search completed: December 16, 2003, 10:52:34

Job time : 50 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2003, 10:52:40 ; Search time 22 Seconds
(without alignments)
1784.749 Million cell updates/sec

Title: US-09-428-122-2
Perfect score: 4774
Sequence: 1 MKSSFPKRVFTFAIFPLM.....FELGSSRNNDLGAKYQF 928

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2802.5	58.7	597	4	US-09-198-452A-29
2	1940	40.6	930	4	US-09-198-452A-470
3	1917.5	40.2	927	4	US-09-198-452A-472
4	1856	38.9	949	4	US-09-198-452A-478
5	1572.5	32.9	643	4	US-09-198-452A-474
6	1417.5	29.7	922	4	US-09-198-452A-15
7	1259	26.4	671	4	US-09-198-452A-468
8	1214	25.4	230	4	US-09-198-452A-30
9	1203.5	25.2	507	4	US-09-198-452A-32
10	1130	23.7	1132	4	US-09-198-452A-466
11	1126.5	23.6	1006	4	US-09-556-877-190
12	1126.5	23.6	1006	4	US-09-620-412C-190
13	1126.5	23.6	1006	4	US-09-598-419-190
14	1124.5	23.6	982	4	US-09-556-877-176
15	1124.5	23.6	982	4	US-09-620-412C-176
16	1124.5	23.6	982	4	US-09-598-419-176
17	1015.5	21.3	450	4	US-09-198-452A-35
18	898	18.6	530	4	US-09-198-452A-482
19	865	18.1	880	4	US-09-556-877-175
20	865	18.1	880	4	US-09-620-412C-175
21	865	18.1	880	4	US-09-598-419-175
22	856	17.9	866	4	US-09-556-877-189
23	856	17.9	866	4	US-09-620-412C-189
24	856	17.9	866	4	US-09-598-419-189
25	834.5	17.5	483	4	US-09-198-452A-27
26	821	17.2	494	4	US-09-198-452A-33
27	792	16.6	427	4	US-09-198-452A-31

28	768	16.1	177	4	US-09-198-452A-28	Sequence 28, Appl
29	708	14.8	294	4	US-09-198-452A-469	Sequence 469, App
30	680.5	14.3	1617	4	US-09-198-452A-1035	Sequence 1035, App
31	667	14.0	1146	4	US-09-198-452A-580	Sequence 580, App
32	646	13.5	969	4	US-09-198-452A-501	Sequence 501, App
33	603.5	12.6	940	4	US-09-198-452A-500	Sequence 500, App
34	582	12.2	1530	4	US-09-556-877-178	Sequence 178, App
35	582	12.2	1530	4	US-09-620-412C-178	Sequence 178, App
36	582	12.2	1530	4	US-09-598-419-178	Sequence 178, App
37	559	11.7	1752	4	US-09-556-877-180	Sequence 180, App
38	559	11.7	1752	4	US-09-620-412C-180	Sequence 180, App
39	559	11.7	1752	4	US-09-598-419-180	Sequence 180, App
40	546.5	11.4	964	4	US-09-556-877-177	Sequence 177, App
41	546.5	11.4	964	4	US-09-620-412C-177	Sequence 177, App
42	546.5	11.4	964	4	US-09-598-419-177	Sequence 177, App
43	546.5	11.4	977	4	US-09-556-877-191	Sequence 191, App
44	546.5	11.4	977	4	US-09-620-412C-191	Sequence 191, App
45	546.5	11.4	977	4	US-09-598-419-191	Sequence 191, App

ALIGNMENTS

RESULT 1
US-09-198-452A-29
; Sequence 29 Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griflais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198.452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 29
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-29

Query Match 58.7%; Score 2802.5; DB 4; Length 597;

Best Local Similarity 93.8%; Pred. No. 5.8e-205; Mismatches 21; Indels 11; Gaps 3;

Matches	560; Conservative	3; Miss	21; In	caps	
QY	149	LSLTNNVSLFKNFSTNDGGAI	TAKTISLTGTTMSALFSEN	TSKKGAIQTSDAL	208
Db	1	LEFDKNVSLFKNFSTNDGGAI	TAKTISLTGTTMSALFSEN	TSKKGAIQTSDAL	60
QY	209	GNQGVFSFSDNTSSDGGAI	FTFASVTISNNAKVSFIDN	KVTGASSTTGDMSG	268
Db	61	GNQGVFSFSDNTSSDGGAI	FTFASVTISNNAKVSFIDN	KVTGASSTTGDMSG	120
QY	269	KTSTDTKVTLTGQMLLF	SNNNTTSTAGGAIYVKLE	ASGGLTIFSRNSVNG	328
Db	121	KTSTDTKVTLTGQMLLF	SNNNTTSTAGGAIYVKLE	ASGGLTIFSRNSVNG	180
QY	329	IAIEDSGELSLSDSGD	IVFLGNTVTSTPGTNRSS	IDLGTSAKMTALRS	388
Db	181	IAIEDSGELSLSDSGD	IVFLGNTVTSTPGTNRSS	IDLGTSAKMTALRS	240
QY	389	PITTSSTVTVDLVKNET	PADSALOYTGNIIFTGKLS	ETEAADSKNLTSLK	448
Db	241	PITTSSTVTVDLVKNET	PADSALOYTGNIIFTGKLS	ETEAADSKNLTSLK	300
QY	449	GGTSLKHGVTLTQFT	QAQADSRLEMDVGTTL	EPADTSTINNLVIN	508
Db	301	GGTSLKHGVTLTQFT	QAQADSRLEMDVGTTL	EPADTSTINNLVIN	360
QY	509	TKATSKNLTSGTITLL	DPTGTFYFNHSLRNP	QSQVDILELKASG	568
Db	361	TKATSKNLTSGTITLL	DPTGTFYFNHSLRNP	QSQVDILELKASG	420

QY 569 PHYGYOCTWGPVWGTGASTATATNTKTYIPNPERIGSLVNSLWNAFIDISSLHLYM 628
 Db 421 PHYGYOCTWGPVWGTGASTATATNTKTYIPNPERIGSLVNSLWNAFIDISSLHLYM 480
 QY 629 ETANEGLOGDRAFWACGLSNFFHKDSTKTRRGFRHLGGVYVIGNLTHTCSDKILSAFQC 688
 Db 481 ETANEGLOGDRAFWACGLSNFFHKDSTKTRRGFRHLGGVYVIGNLTHTCSDKILSAFQC 540
 QY 689 LFGDRDYFVAKNGQ-TVYGGTLYYQHNETHYIS-----LPCKLR-PCSLSYYP 734
 Db 541 LFGDRDYFVAKNGQYSLRRNSLLPAQNLNLYLSLQTTALFVVLVCSYRSCSLFRKP 597

RESULT 2
 US-09-198-452A-470
 ; Sequence 470, Application US/09198452A
 ; Patent No. 6559294
 ; GENERAL INFORMATION:
 ; APPLICANT: Griffais, R.
 ; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
 ; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
 ; TITLE OF INVENTION: and treatment of infection
 ; FILE REFERENCE: 9710-003-999
 ; CURRENT APPLICATION NUMBER: US/09/198,452A
 ; CURRENT FILING DATE: 1998-11-24
 ; NUMBER OF SEQ ID NOS: 6849
 ; SEQ ID NO 470
 ; LENGTH: 930
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 US-09-198-452A-470

Query Match 40.6%; Score 1940; DB 4; Length 930;
 Best Local Similarity 43.7%; Pred. No. 5e-139;
 Matches 416; Conservative 165; Mismatches 324; Indels 48; Gaps 17;

QY 1 MKSSFPKVFSTPAFLPLSM-IAT---ETVLDSSASFDGKNGNFSVRESQEDAGTYL 56
 Db 1 MKIPLKLLISTLVTILLIATYAGDASLPTSDFGAGGSTFTPKSTADANGTYNL 60
 QY 57 KGNVTLENI PGCTGTAITKSCFNPNKGDLTFTGNGNSLLFQTVADGTVAGAAVNSVVVDK 116
 Db 61 SGNVYI-NDACKGALTCCCTETETGDLTFTGKGYSFNTVDAGSNAGAAA-STADKA 118
 QY 117 TTFGFSLSPIASPGSIITKGAIVSCSTGSLSTKXNLLFSKNFSTD---NGGAITA 173
 Db 119 LTFTGFSNLSPIAAGPTTVASGKSTLS-SAGALNTDNGTILFSQNVSEANNNGGAITA 177
 QY 174 KTLSTGTMTGALFSENTSSKKGAIQTSALTITGNGEVSFSDNTSSDSGAAIFTEAS 233
 Db 178 KTLISGNTSSITITSSAKKLGALVSSAAASIGNTGQLVFNNGETGGGALGFEAS 237
 QY 234 VTISNAKVSIIDKNTGASSTTGDMSGGGAICAYKTSPTKVTLTQNMQLFSNNSTT 293
 Db 238 SSITQNSLFFSGNTATDAAG-----KGGAIYCEKGETPTLTISGNKSLTFAENSVT 291
 QY 294 AGGAIYVKLELASGLTLFERNVNGTAPKGAIAIEDSGELSLSDSGDIVFLGNTV 353
 Db 292 QGGALCAHGLDLSAAGPTLTFNNRCGNTAAGKGAIAIADSGSLSANQGDITFLGNTL 351
 QY 354 TSTT-PGTVNRSSIDLTSAKWTALRSAGRAIYVDPI---TTGSSYTVTDVLKVNTPA 409
 Db 352 TSTSAPTSTRAIYLGSSAKITNLRAAQOQSIYVDPIASNTTCAS---DVLTIQPD 407
 QY 410 DSALQYTNITPTGKLBETBAASKNLTSKLPVLSGGLSLKKGVTLQTAQFOQA 469
 Db 408 NSPLDYSGTIVFSGEKLGADEAKADNFTSILKQPLALASGTLAKGNVELDVNGPTQTE 467
 QY 470 DSRLEMDVGTILEPADTSTIN--NLVINISIDGAKKAKIETKATSKNLTLSGHTILLDP 527
 Db 468 GSTLLMQPGTKLK-ADTEAISLTKLVDLSALEGNKSVSIETAGANKTITLTSPLVQDS 526

QY 528 TGTTFVNHSLRNPOSYDILELKAS-----GTVTSTAVTTPDIPMGERFHYGQCTWG 578
 Db 527 SGNVESHTINCAFTQPLVVFATAAASDIYIDALLTSPVQTPEP-----HYGYQHW 580
 QY 579 PIVNGTGASTTATFNWTKTYIPNPERIGSLVNSLWNAFIDISSLHLYMETANEGLOGD 638
 Db 581 ATWADTSTAKSGITWTTTGYNPNPERRASVVPDSLWASFTDRTLQQIMTSQANSIYQ 640
 QY 639 RAFPCAGLSNFFHKDSTKTRRGFRHLGGVYVIGNLTHTCSDKILSAFQCFGRDRDYFV 698
 Db 641 RGLWASGTANFFHKDSTKTRRGFRHLGGVYVIGNLTHTCSDKILSAFQCFGRDRDYFV 700
 QY 699 AKNQCTVYGGTLYYQHNETHYISLPCKLKRPC--SLSYVPTPIPVLFSGNLSYTHTDNDLKT 756
 Db 701 VENTSHNYLASLYLQHRAFLGGLP--MPSFGSITDMLKDIPILNLAQLSYSTKNDMDT 757
 QY 757 KYTTYPTVKGWGNDSFALEFGGRAPICL-DESALFQYMPFMKLPVYAHQGFKEQGT 815
 Db 758 RYTSYPEAOGSWTNNSGALELGGSLALYLPKEAPFFQGYFFFLKFPQAVYRQONFKESGA 817
 QY 816 EAREFGSSRLVNLALPIGIRPDKESDQDATYNLTGLTVDLVRSNPDCTTLIRISGDSW 875
 Db 818 EARAFTDDGDIYVNCISIPVGIKLEKISEDEKNNFEISLAYIGDVYKKNPRSTSLMVSGASW 877
 QY 876 KTFGTNLARQALVLRAGNHCNFSNFAFSQFSFELRGSSRNYNVDLGAKYQF 928
 Db 878 TSLCKNLARQAFSLASAGSHLTLSPHVELSGEAAVELRGSAAHYNVDCGLYSF 930

RESULT 3
 US-09-198-452A-472
 ; Sequence 472, Application US/09198452A
 ; Patent No. 6559294
 ; GENERAL INFORMATION:
 ; APPLICANT: Griffais, R.
 ; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragme
 ; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, pr
 ; TITLE OF INVENTION: and treatment of infection
 ; FILE REFERENCE: 9710-003-999
 ; CURRENT APPLICATION NUMBER: US/09/198,452A
 ; CURRENT FILING DATE: 1998-11-24
 ; NUMBER OF SEQ ID NOS: 6849
 ; SEQ ID NO 472
 ; LENGTH: 927
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 ; NAME/KEY: SITE
 ; LOCATION: 1...927
 ; OTHER INFORMATION: Xaa=unknown or other
 US-09-198-452A-472

Query Match 40.2%; Score 1917.5; DB 4; Length 927;
 Best Local Similarity 44.8%; Pred. No. 2.6e-137;
 Matches 427; Conservative 147; Mismatches 328; Indels 51; Gaps 22;

QY 1 MKSSFPKVFSTPAFLPLSM-----IATVLDSSASFDGKNGNFSVRESQEDA-GTT 53
 Db 1 MKSSLHWFLISSLSLALPLSLNFSAPAUVVEINLGPNTSFG--PCTYTPPAQTNAQGTI 58
 QY 54 YLFGKNTLENI PGCTGTAITKSCFNNTKGDLTFTGNGNSLLFQTVADGTVAGAAVNSV 113
 Db 59 YNLTGDVSI TN-AGSPALTATACFPKETTGNLSFGHGQYQLQNI DAG--ANCTFTTAA 115
 QY 114 DKSTTFIGFESSLFTASPGSSITTKGAVSCSTGSLSTKXNLLFSKNFSTONGGAITA 173
 Db 116 NKLLSFGSFSYLSLQI--TNATGTGAIK-STGACSIQSNYSYFCQNFNDNGGALQG 172
 QY 174 KTLSTGTMTSALFSENTSSKKGAIQTSALTITGNGEVSFSDNTSSDSGAAIFTEAS 233
 Db 173 SSISLS-LNPULTFAKNKATQKGLYSGTIGTINTNLNSAFSENTAANGGAIYFAS 231
 QY 234 VTISNAKVSIIDKNTGASSTTGDMSGGGAICAYKTSPTKVTLTQNMQLFSNNSTT 292

Db 232 SFISSNKAISFINNSVTATSA-----TGGAIYCSSTAPKPVLTSLDNGELNFIGNTAI 285
Qy TAGGAIYVVKLELASGLTLPFRNSVNGGTAPKGAIAIEDSGELSLSDSGDVFLENGT 352
Db TSGGAIYTDNLVLSGGPTELPKNSAIDTAAPLGAIAIADSGSLSLALGGDITFEGNT 345
Qy V---TSTPTGTRNSIDIG-TSAKWTALRSAGRAIYFYDPTTGGSTTVTDVLKVNETP 408
Db VVKGASSSQTTTRNSINIGNINAKIVQLRASQGNITFYDPTTISITAAALSADALNAGPD 405
Qy ADSALQYTGNIIFTGCKLSETEADSKULTSKLQPVTLTSGTSLXKGVTLQCAFTQQ 468
Db LAGNPAYQGTIVFSGEKLSEAAEADNLKSTIOOPLTAGGQJLSKSGVTLVAKSFQS 465
Qy ADSLEMDVGTLEPADSTINNLVINISIDGAKKAIET-KAT--SKNLTSLTGTITLL 525
Db PGSTLLMDAGTTLTADGS-----LSIICSQRFKRDQEXTLKATQASQTVTLSGSLV 521
Qy DPTGTTFYENHSLRNPQSYDIIELKAS--GTVTSTAVTPDPINGEKPHFYQGTWGPVWG 583
Db DPGNVIEDVSNWNPQVFSCLTTLTADDPANIHTDLAADPLEKNPJHGWYQGNWA-LSWQ 580
Qy TGAST---TATFNWTKTGYPNPERIGSLVPSNLWNAFIDISLHLMETANEGLQGDRA 640
Db EDTATKSAATLTWTKTGNPNPERRGTLVANTLWGSFVDVRSIQOOLVATKVRSQETRG 640
Qy FWCAGLSNPFHDKSTKTRGPHLGGYVIGNLHTCSDKILSAAPCOLFGDRDRDYFAVK 700
Db IWCBSNPFHDKSTKINKGPHISAGYVVGATTTLASDNLITAAFCOLFGKDRDHFINK 700
Qy NGQTVYGGTLYYQHNETYISLPCRLPCSLSVYP---TEIPVLPFGNLSYTHTDNDLTKT 757
Db NRASAYASLHLQHLATLSS-----PSLLRYLPGSESEQPVLFDAQISYIYXKNTMKTY 754
Qy YTYPTVKGWNGDSFALBFGGRAP-ICLDESALFEQYMPFMKLOFVYAHQGEKEQGT 816
Db YTAQPKGESSWYNDGCALELASSLPHALSHEGLFHAYFPFIKVEASYIHQDSFERNTT 814
Qy -ABRGSSRLNALPIGIRFDEKESCDATYNTLGYTVDLVRSNPDCTTTLRISGDSW 875
Db LVRSFDSGDLINVSPIGITEFRFSNERASTVATVIYADVVRKNPDCTTALLNNTSW 874
Qy KTFGTNLARQALVLRAGNEHCFNSNPEAFSQFSFELRGSSRNYNVDLGAKYQF 928
Db KTGITNLSRQAGIGRAGIFVAFSPNLEVTNLSMEIRGSSRSYNADLGCKFQF 927

RESULT 4
US-09-198-452A-478
; Sequence 478, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198, 452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 478
; LENGTH: 949
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 1...949
; OTHER INFORMATION: Xaa=unknown or other
US-09-198-452A-478

Query Match 38.9%; Score 1856; DB 4; Length 949;
Best Local Similarity 42.7%; Pred. No. 1.3e-132;

Matches 405; Conservative 171; Mismatches 331; Indels 42; Gaps 20;
Qy 1 MKSFPKFPVSTAIIP--LSMLATEFVLDSASFQGN-KNGNFSVRESQEDAGTTVLFK 57
Db MKTSIPVWLVSFLAFSCHLOSLANEELSPDOSFNGNIDSGTFPKTS-----ATTSLT 77
Qy 58 GNVLTENIPGTGTAITKSCFNNTKGLTFTTCNGNSLLFQTVDAQTAVAGAAVNSVDKST 117
Db 78 GDVFFYE-PKQGPLSDSCPKQITDNLTLFLNGHSLTFGFDAGTHAGAAA-STANKNL 135
Qy 118 TFGFSSLSFIAPSGSSITTKGAVSCSTGSLSTKUNVLSLFSKNFSTGNGGAIKATLS 177
Db 136 TFGFSLSPDSSTVTTTQGTLS-SAGGVNLENKLVVAGNFSTADGAGAKGASFL 194
Qy 178 LTGTTMALPSENTSSAKGGAIOQSDALTTTGNQGEVSFSDNTSSDGAATFTEASVTIS 237
Db 195 LTGTSGDALFSSNSSSTKGGAIATTAGARIANNKTVRFLSNIASTSGGAIDDGTSILS 254
Qy 238 NNAKVSFIDNKTVCASSSTTTCGMSGGAICAYKTSDDTKVTLTGQMLLFSNNTTATAGA 297
Db 255 NKKLYP-----EGNAAKTT---GGALCNTKASGSPELIISNNKTLIFASNVAETSGA 305
Qy 298 IYVKKLELAGGLTLFGRNSVNGGTAPKGAIAIEDSGELSLSDSGDIVPLGNVTIST- 356
Db 306 IHAKKALSSGGTFEFLRNNVSSAT-PKGGAIISIDASGELSLSAETGNITFVRNTLTGT 364
Qy 357 -TPGTRSSIDLGTSAKWTALRSAGRAIYFYDPTTGGSTTVTDVLKVNTPADSAQY 415
Db 365 STDTPKRNAINIGSNGKFEFLRAAKNHTIFYDPTT---SEGTSSDVLKINNGSAGALNPY 422
Qy 416 TGNITFTCEKLSSETAADSCKNLTSKLQPVTLTSGTSLXKGVTLQOAFQOADSRLM 475
Db 423 QGTILFSGETLTADKLVADNLKSSFTQVSLSGKLLAQKGVTLSESTFSQEGAGSLGM 482
Qy 476 DVGTTLE-PATSTINNLVINISSIDGAKKAIETAKSKNLTSLGTTILDPTGTFYEN 534
Db 483 DSGTTLTSTAGTITITNLGINVDLSGLKQPVSLTAKGASNKVIVSGKLNLDIEGNIYES 542
Qy 535 HSLRNPQSYDIIELKASGTVTS---TAVTPDPINGEKPHFYQGTWGPVWGTGASTT- 589
Db 543 HMFSDQLFSLKLTVDADVDVTNVDISLIPVAEDPNSEYFGQGNV-VNWTDTATNT 601
Qy 590 --ATENWTKGTGYPNPERIGSLVPSNLWNAFIDISLHLMETANEGLQGDRAFCAGLS 647
Db 602 KEATATWTKTGTFVSPERKSAVLCNLTWGVFTDLSLQOLVIGATGMEHKQGFVSSNT 661
Qy 648 NPFHKDSTKTRGPHLGGYVIGNLHTCSDKILSAAPCOLFGDRDRDYFVAKNOGTVYG 707
Db 662 NFLHKTGDNKRKGRHTSGGYVIGGSAHTPKDDLFTFAFCHLFARDKDCFIAHNNSRTYG 721
Qy 708 GILYQHNET-----YISL-PCKLRPCSLSYVPTPIPLVPSGNLSYTHTDNDLTKYTY 761
Db 722 GTLPFKHSHTLOPQNYLRGRAKFSESAIEKFPREIPALDVQVFSFSDNRMETHYISL 781
Qy 762 PTVKSGWNGDSFALBFGGRAPICL-DESALFEQYMPFMKLOFVYAHQGEKEQGTAREF 820
Db 782 PSESGWSNECLAGGILDLPLFVLSNPFLPFTFIPQMKVENVYVSQNSFFSSSDGRGF 841
Qy 821 GSSRLVNLALPIGIRFDEKESDQDA-TYNLTIGYTVDLVRSNPDCTTTLRISGDSWKTG 879
Db 842 SIGRLNLISIPVCAKF-VQGDIGDSYTYDLSGFFVSDVYRNNPQSTATLVMSPDSSWKIRG 900
Qy 880 TNLARQALVLRAGNEHCFNSNPEAFSQFSFELRGSSRNYNVDLGAKYQF 928
Db 901 GNLSRQAFLLRGSNNYVYNSNCELFGHYAMELRGSSRNYNVDVGTGLRF 949

RESULT 5
US-09-198-452A-474
; Sequence 474, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.

Wed Dec 17 09:17:09 2003

; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 474
; LENGTH: 643
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; US-09-198-452A-474

Query Match 32.9%; Score 1572.5; DB 4; Length 643;
Best Local Similarity 49.8%; Pred. No. 2.6e-111;
Matches 314; Conservative 93; Mismatches 209; Indels 15; Gaps 6;

QY 308 GGLTFLSRNSVNGGTPAPKGAIAIEDSGELSLSDSADSDIVFLGNTVSTTP-GTNRSSID 366
D 18 GGGISFNNIVQTTAGNGGALSILAAGCSLSAEAGDITFNGNAIVATTPQTKRNSID 77
QY 367 LGTSAXMTALRSAGRAIFYDPITTTGSSSTTVTVLKNVETPADSALQYTGNIIFTGKEL 426
D 78 IGSTAKITNLRLSGHSIFPYDIPITANTAADSTDLNLKADAGNSTDYSGSIVFSGKEL 137
QY 427 SETEAADSKNLTKSLKLPVTLGGTSLKHGVTLOTQAFQQQADSLMDVGTTLPEA-D 485
D 138 SEDEAKVADNLSTLQPVTLTGNLVLRGVTLDTKGFTQTAGSSVIMDAGTTLKASTE 197
QY 486 TSTINNLVINISIDGAKKAKIETKATSKNLTKSLGNTLLDPTGFVENHSLRNPQSYDI 545
D 198 EVLTGLTSLPVDLGBGKVVIAAASAKSNVALSGPILLDDNQNAYENDHDKTQDFSF 257
QY 546 LELKASGTVTSTAVTPDPIPMGEKPHYGYQGTWGPVW-----GTGASTTATFNNWTKGYI 600
D 258 VQSLGATATTVDVPAVPTVATPHYGYQGTW-G-MTWVDDTASTPKTKTATLAWTNTGYL 316
QY 601 PNERIGSLVPSLWNAFIDISLHYLMETANEGLQGDRAFCAGLSNPFHKDSTKTRG 660
D 317 PNERQGPLVNSLWGSFSDIQAIOGVIERALSALTLCSDRGFWAAGVANFLDKDKKGEKRK 376
QY 661 FRLHSGGYVIGNLTCSDKLSAFCOLGRDREDFVAKNQCVYGGTYLYOHNERYIS 720
D 377 YRHSGGYAIGGAQTCSENLISAFQOLGSDKDFLAKNHTDTYAGAFYIQL-----IT 432
QY 721 LPCKLRPCSLSYVP---TEIPVPSGNLSYTHTDNLTKYTTYPTVKGSGWGNDSFALEF 777
D 433 ECSGFICGLLDKLPGSWSHKLVLLEGOLAYSHVSNLDLTKYTAYPEVKSGWGNNAFNML 492
QY 778 GGRAPICLDESALPEQYMPFMKLOFVYAHQBGFKQGTAEERFSGSSRLVNLALBPIGRFD 837
D 493 GASSHSYPEYLHCFDTPAYIKLNLTYIRODSFSEKGTGEGRSFDDSNLNLSLPIGVKFE 552
QY 838 KESDCODATYNLTGYTVDLVRSNPDCTTLIRISGDSWKTFTGLARQALVLRAGNHFCF 897
D 553 KFSPCNDFSYDLTSLVSPDLIRNPKCTTALVISGASWETVANNLARQALQVRAGSHYAF 612
QY 898 NSNFEAFQGFELRGSSRNYNVDLGAQYQF 928
D 613 SPMEVLQGFVEVRGSSRIYNVDLGGKFQF 643

RESULT 6
US-09-198-452A-15
; Sequence 15, Application US/09198452A
; Patent No. 659294
; GENERAL INFORMATION:
; APPLICANT: Grifais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A

; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 15
; LENGTH: 922
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 1...922
; OTHER INFORMATION: Xaa=unknown or other
; US-09-198-452A-15

Query Match 29.7%; Score 1417.5; DB 4; Length 922;
Best Local Similarity 36.5%; Pred. No. 2.9e-99;
Matches 347; Conservative 156; Mismatches 385; Indels 63; Gaps 26;

QY 5 PP-KPVFTFAIFPLSMIATETVLDSSASFDQNKGNFVSRESQBEDAGTTLFKGNVTLE 63
D 8 PFLVFSFTLLSVFDTLSGATTISLTPEDSFHGDSQN--AERSYNVQAGDVVSLTGDVSLIS 65
QY 64 NIPGTGTAITKSCNNTYKDLTFTGNGNSLLPQTVDAAGTVAGAAVSSVVDKSTT-FIGF 122
D 66 NV--DNSALNKACFVTVSGSVTFAGHHGXYPNNISSGTTKEGAVLCCODPQATARFSGF 123
QY 123 SLSLFTASPGSSITTKGAVSC--STGSLSLTKNVSLLSFKNFSNDNGAIFTAKTLSLTG 180
D 124 STLSPNQSPGDIKEQ-----CLYSKNALMLLNNVVRPEQKQSKTKGGAISGANVTIVG 178
QY 181 TTMSALFSENSTSKKGGAIOQTSDALITIGNQGEVFSFSDNTSSD-SCAAIFTASVTSNN 239
D 179 NYDSVSFYQNAAT-FGGAIHSSGPIQIAVNOAIRFAQNTAKNGSGGALYSGDDIDIDQN 237
QY 240 AKVSFIDNKTVCASSSTTCDMSGCAICAYKTSTDTK---VTLTGNQMLLFSNNTSTTAG 295
D 238 AVLVFENE-----AULTAIGKGAVCCLPSTGSSSTFPVIVTFSDNKQLVFRNHSIMG 292
QY 296 GAIYVKKLBLASGGLTFLSRNSVNGGTPAPKGAIAIEDSGELSLSDSADSDIVFLGNTVTS 355
D 293 GAIYARKLSISSGGPTLFINNISYANSQNLGGAIAIDTGEISLSAEKGTITFQGN--RT 350
QY 356 TTPGTNRSSIDLTSAXMTALRSAGRAIFYDPITTTGSSSTTVTVLKNVETPADSALQY 415
D 351 SLSPFLN-GIHLHLQNAKFLKLQARNGYSIFDYDPIIT--SEADGSTQLNINGDPKNK--EY 404
QY 416 TGNIIPTGKLSLSTEADSNLTKSLKLPVTLGGTSLKHGVTLOTQAFQQADSRLEM 475
D 405 TGIILFSGEK--SLANDPRDPKSTIPQNVNLASGLVIVKEGAETVTSKTSQSPGSHVL 461
QY 476 DVGTTL-EPADTSTINNLVINISIDGAKKAK-IETKATSKNLTKSLGNTLLDPTGTFYE 533
D 462 DLGKTLFASKEDIAITGLAIDIDSLSSSSTAATAVIKANTANKQISVTDSELSISPTGNAYE 521
QY 534 NLSLRNPQSDVILELK--ASGTVTSTA-----VTPDPIMGEKPHYGYQGTWGPVW-GTG 585
D 522 DLRRNSQTPFLLSLEPAGAGSVTVTAGDFLPVSP-----HYGFQGNW-KLAWTGTG 572
QY 586 ASATTATNTWTKGYIPNPERIGSLVPSLWNAFIDISLHYLMETANEGLQGDRAFCAG 645
D 573 -NKVGGEFWDKINIKPRPEKEGNLVFNLLMGNAVDSLMQVQVETHASSLQTDRLWIDG 631
QY 646 LSNFFHKDSTKTRGRFRLHSGGYVIGNLTCSDKLSAFCQLFGDRDRDYFAKNGQTV 705
D 632 IGNLFHVSASEDNIRYRHNSGGYVLSVNNETTPKHYSMAFSQLSFSDKDYAVSNNEIRM 691
QY 706 YGGTYLYOHNET-----YISLPCKLRPCSLSYVPTPIP-VLFSGNLSYTHTDNLTKY 758
D 692 YLGSYLYQYTTSLGNIIFRYASRNPNVNVGILSRFLQNLPLMIFHLCAIGHATNMDKTDY 751
QY 759 TTYPTVKGSGWGNDSFALEFGGRAPICLDESA-LFEQYMPFMKLOFVYAHQBGFKQGTGTEA 817
D 752 ANFPVKNWNRNWCALIECGSGMELLVPENGRLFQGAIPFMKLOLVYAYQGFKEITADG 811
QY 818 REFSSRLVNLALPIGIRFDKESDQDATYNLTGYTVDLVRSNPDCTTLIRISGDSWKT 877

841 SDVRWNNCLAGEITAGLPVITPSKLYLNLRFVQAEFSYADHESFTEEGDQARAFKS 900
823 SRLVNLALPTGIRDFKESDCQDATYNTLTVLVRNSNPDCTTTLIRISGDSWITFGTNL 882
901 GHLNLVSVGVKDFRCSSHPKYSFMAAYICDAYRTISGTETTLSSHQETWTTDAFHL 960
883 ARQALVLRAGNHFCFNSFPAFQSFELRGSSNNYVNDLGAKYQF 928
961 ARHGVVVRGSMYASLTNIEVYGRVYRDASRGYGLSAGSKVRF 1006

RESULT 12
US-09-620-412C-190
Query Match 23.6%; Score 1126.5; DB 4; Length 1006;
Best Local Similarity 30.1%; Pred. No. 4.6e-77;
Matches 303; Conservative 160; Mismatches 372; Indels 171; Gaps 30;
GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 190
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-620-412C-190

Query Match 23.6%; Score 1126.5; DB 4; Length 1006;
Best Local Similarity 30.1%; Pred. No. 4.6e-77;
Matches 303; Conservative 160; Mismatches 372; Indels 171; Gaps 30;
GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 190
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-620-412C-190

841 SDVRWNNCLAGEITAGLPVITPSKLYLNLRFVQAEFSYADHESFTEEGDQARAFKS 900
823 SRLVNLALPTGIRDFKESDCQDATYNTLTVLVRNSNPDCTTTLIRISGDSWITFGTNL 882
901 GHLNLVSVGVKDFRCSSHPKYSFMAAYICDAYRTISGTETTLSSHQETWTTDAFHL 960
883 ARQALVLRAGNHFCFNSFPAFQSFELRGSSNNYVNDLGAKYQF 928
961 ARHGVVVRGSMYASLTNIEVYGRVYRDASRGYGLSAGSKVRF 1006

RESULT 12
US-09-620-412C-190
Query Match 23.6%; Score 1126.5; DB 4; Length 1006;
Best Local Similarity 30.1%; Pred. No. 4.6e-77;
Matches 303; Conservative 160; Mismatches 372; Indels 171; Gaps 30;
GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 190
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-556-877-190

Query Match 23.6%; Score 1126.5; DB 4; Length 1006;
Best Local Similarity 30.1%; Pred. No. 4.6e-77;
Matches 303; Conservative 160; Mismatches 372; Indels 171; Gaps 30;
GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 190
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-556-877-190

Db 597 DRYDWLGSNQKINVLKQL-GTKPPANAPSDLTGNEMPKYGYQGSW-KLAWDPNTANN 654
QY 591 TP----NWTKTGYIPNPERIGSLVPSNLWNAFIDISSLHYLMETANEGLQGDRAFWCAGL 646
Db 655 PYTLKATWTKTGYNPGFERVASLVPSNLWGSILDIRSAHSAIQASVDGRSYCRGLWWSGV 714
QY 647 SNFFHXDSTKTRRGRHLSGGYVIGGNLHTCSDKILSAAFCOLFCGRDRDYFVAKNQGTVY 706
Db 715 SNFFYHDDRDLGQGYRYISGGYSLGANSYFGS-SMEGLAFTVEVFGSKDYVVCRSNHHAC 773
QY 707 GGTLYYQHNETYISLPCKLRPCSLSYVPTPEIPVLPFG---NLSYTHDNDLTKYTYPT 763
Db 774 IGSVYLSQQAL-----CG-SY-----LFGDAFIRASYGFGNQHMKTSYTPAE 816
QY 764 VKSGWGNDSFALEFGGRAPICLDESALP-EQYMPFMKIQVYAHQEGFKEQGTAEAREFGS 822
Db 817 SDVRWNNCLAGEIGAGLPIVITPSKLYLNLRLPFVQAEFSYADHESFTEEGDQARAFKS 876
QY 823 SRLVNLALFIGIRPKESDQDATYNLTGTVDLVRSNPDCTTTLIRISGDSWKTFTGNL 882
Db 877 GHLLNLSVPVGVKFDRCSSHPNKYSFMAAYICDAYRTISGTETTLSSHQETWTTDAFHL 936
QY 883 ARQALVLRAGNHFCEFSNFEAFSQFSFELRGSSRNYNVDLGAKYOF 928
Db 937 ARGVVVRGSMYASLTSNIEVYGHGYEYRDASRGYGLSAGSKVXP 982

Search completed: December 16, 2003, 10:55:18
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2003, 10:54:06 ; Search time 37 Seconds
(without alignments)
4664.671 Million cell updates/sec

Title: US-09-428-122-2
Perfect score: 4774
Sequence: 1 MKSFPKVFSTFAIFPLSM.....FELRGSSRYNVDLGAQYQF 928

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4774	100.0	928	12	US-09-428-122-2
2	1965	41.2	936	9	US-09-452-380-3
3	1965	41.2	936	12	US-10-324-129-3
4	1946.5	40.8	925	9	US-09-452-380-4
5	1946.5	40.8	925	12	US-10-324-129-4
6	1915	40.1	926	12	US-09-738-269-57
7	1915	40.1	926	14	US-10-023-437-57
8	1656.5	34.7	839	12	US-09-738-269-23
9	1656.5	34.7	839	14	US-10-023-437-23
10	1437.5	30.1	922	9	US-09-886-468-19
11	1126.5	23.6	1006	9	US-09-841-132-190
12	1124.5	23.6	882	9	US-09-841-132-176
13	865	18.1	880	9	US-09-841-132-175
14	856	17.9	866	9	US-09-841-132-189
15	851.5	17.8	679	12	US-09-738-269-53

16	851.5	17.8	679	14	US-10-023-437-53	Sequence 53, Appl
17	737.5	15.2	439	9	US-09-841-132-524	Sequence 524, App
18	685.5	14.4	931	10	US-09-779-081-2	Sequence 2, Appli
19	685.5	14.4	1609	12	US-09-841-260-75	Sequence 75, Appl
20	685.5	14.4	1609	14	US-10-007-693-75	Sequence 75, Appl
21	672.5	14.1	978	12	US-09-841-260-65	Sequence 65, Appl
22	672.5	14.1	978	14	US-10-007-693-65	Sequence 65, Appl
23	666	14.0	1723	9	US-09-841-132-394	Sequence 394, App
24	666	14.0	1723	9	US-09-841-132-395	Sequence 395, App
25	608	12.7	871	9	US-09-886-468-21	Sequence 21, Appl
26	597	12.5	1016	12	US-09-841-260-95	Sequence 95, Appl
27	597	12.5	1016	14	US-10-007-693-95	Sequence 95, Appl
28	582	12.2	1530	9	US-09-841-132-178	Sequence 178, App
29	580.5	12.2	963	9	US-09-886-468-22	Sequence 22, Appl
30	576.5	12.1	1531	12	US-09-841-260-98	Sequence 98, Appl
31	576.5	12.1	1531	14	US-10-007-693-98	Sequence 98, Appl
32	559	11.7	1751	9	US-09-841-132-445	Sequence 445, App
33	559	11.7	1751	9	US-09-841-132-594	Sequence 594, App
34	559	11.7	1752	9	US-09-841-132-180	Sequence 180, App
35	546.5	11.4	964	9	US-09-841-132-177	Sequence 177, App
36	546.5	11.4	977	9	US-09-841-132-191	Sequence 191, App
37	523	11.0	1770	9	US-09-841-132-444	Sequence 444, App
38	513	10.7	1776	9	US-09-841-132-179	Sequence 179, App
39	502	10.5	848	9	US-09-841-132-192	Sequence 192, App
40	490.5	10.3	948	9	US-09-841-132-194	Sequence 194, App
41	477.5	10.0	514	9	US-09-886-468-23	Sequence 23, Appl
42	394	8.3	691	9	US-09-841-132-313	Sequence 313, App
43	385	8.1	162	12	US-09-738-269-55	Sequence 55, Appl
44	385	8.1	162	14	US-10-023-437-55	Sequence 55, Appl
45	344	7.2	700	9	US-09-841-132-345	Sequence 345, App

ALIGNMENTS

RESULT 1

US-09-428-122-2
; Sequence 2, Application US/09428122
; Publication No. US20030170259A1
; GENERAL INFORMATION:
; APPLICANT: Connaught Laboratories Limited
; APPLICANT: Murdin et al.
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 19721-007-019
; CURRENT APPLICATION NUMBER: US/09/428,122
; CURRENT FILING DATE: 1999-10-27
; EARLIER FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: 60/106,046
; EARLIER FILING DATE: 1999-05-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 928
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; US-09-428-122-2

Query Match	100.0%;	Score 4774;	DB 12;	Length 928;
Best Local Similarity	100.0%;	Pred. NO. 0;		
Matches 928;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MKSFPKVFSTFAIFPLSMIATETVLDSSASFDGKNGNFSVRSQEDAGTTLFGKGV	60	
Db	1	MKSFPKVFSTFAIFPLSMIATETVLDSSASFDGKNGNFSVRSQEDAGTTLFGKGV	60	
QY	61	TLENIGTGTAITKSCFNNTKGLTFTGNLSLLFOTVDAGTVAGAAVNSVVDKSTTFI	120	
Db	61	TLENIGTGTAITKSCFNNTKGLTFTGNLSLLFOTVDAGTVAGAAVNSVVDKSTTFI	120	
QY	121	GFSSLSFIASPGSSITTKGAVSCSTGSLTKNLSLLFSKNFSTDNGGAITAKTISLTG	180	

Db 121 GFSSLSFIASPGSSITTTGKGVSCSTGSLSTKVNLSLFFSKNFSTNDGGAITAKTLTLTG 180
Qy 181 TTMSALFSENTSSKKGGAIOQSDALTITGNQGEVSFSDNTSSDGAALFTEASVTISNNA 240
Db 181 TTMSALFSENTSSKKGGAIOQSDALTITGNQGEVSFSDNTSSDGAALFTEASVTISNNA 240
Qy 241 KVSFIDNKVTCASSTTGDMSGGAICAYKTSDDTKVTLTGQNMLLFSNNTSTTAGGAIYV 300
Db 241 KVSFIDNKVTCASSTTGDMSGGAICAYKTSDDTKVTLTGQNMLLFSNNTSTTAGGAIYV 300
Qy 301 KKELEASGLTLFERNVNGTAPKGAIAIEDSGELSLSDSDGDIIVFLGNVTSSTTPTG 360
Db 301 KKELEASGLTLFERNVNGTAPKGAIAIEDSGELSLSDSDGDIIVFLGNVTSSTTPTG 360
Qy 361 NRSSIDLGTSAKMTALRAAGRAIYFYDPIITGSSSTTVDLKYNETPADSALQYTGNI 420
Db 361 NRSSIDLGTSAKMTALRAAGRAIYFYDPIITGSSSTTVDLKYNETPADSALQYTGNI 420
Qy 421 FTGEKLSSETAADSNNLSKLLQPVTLSSGGLSLKHGVTLOTQFTQOADSRLMDVGT 480
Db 421 FTGEKLSSETAADSNNLSKLLQPVTLSSGGLSLKHGVTLOTQFTQOADSRLMDVGT 480
Qy 481 LEPADTSIINNVINISIDGAKAKIETKATSKNLTLSGTTITLDDPTGTFYENHSLRNP 540
Db 481 LEPADTSIINNVINISIDGAKAKIETKATSKNLTLSGTTITLDDPTGTFYENHSLRNP 540
Qy 541 QSYDILELKASGTVTSTAVTPDPIMGEKPHYGYQGTGPIVWGTGASTTATFNWTKTYI 600
Db 541 QSYDILELKASGTVTSTAVTPDPIMGEKPHYGYQGTGPIVWGTGASTTATFNWTKTYI 600
Qy 601 NPFRIGSLVNSLWNAIDISSLHYLMETANEGLQDRAFWCAGLSNFHKSSTKTRRG 660
Db 601 NPFRIGSLVNSLWNAIDISSLHYLMETANEGLQDRAFWCAGLSNFHKSSTKTRRG 660
Qy 661 FHLSSGVVIGNLTSCDKILSAFQCFGRDRDYFVAKNQGTGTYGTYQHNETHYIS 720
Db 661 FHLSSGVVIGNLTSCDKILSAFQCFGRDRDYFVAKNQGTGTYGTYQHNETHYIS 720
Qy 721 LPCKLRPCSLSVPTTEIIVLFSNLSYTHNDNLKTKYTTTPTVKGSGWNSDPALEFGGR 780
Db 721 LPCKLRPCSLSVPTTEIIVLFSNLSYTHNDNLKTKYTTTPTVKGSGWNSDPALEFGGR 780
Qy 781 APICLDESALPEQYMFPMKLOFVVAHQGFKEQGTAREFGSSRLVNLALPIGRFRDKES 840
Db 781 APICLDESALPEQYMFPMKLOFVVAHQGFKEQGTAREFGSSRLVNLALPIGRFRDKES 840
Qy 841 DCQDATYNLTGYTVDLVRSNPDCTTLIRISGDSWKTFTGTMARQALVLRAGNHFCFNSN 900
Db 841 DCQDATYNLTGYTVDLVRSNPDCTTLIRISGDSWKTFTGTMARQALVLRAGNHFCFNSN 900
Qy 901 FFAFSQFSFELRGSSRNVDLGAKYQF 928
Db 901 FFAFSQFSFELRGSSRNVDLGAKYQF 928

RESULT 2

US-09-452-380-3
; Sequence 3, Application US/09452380
; Patent No. US20020094340A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, Andrew D.
; APPLICANT: COHEN, Raymond P.
; APPLICANT: WANG, Joe
; APPLICANT: DUNN, Pamela
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE REFERENCE: 032931/0216
; CURRENT APPLICATION NUMBER: US/09/452,380
; PRIOR FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: US 60/132,272
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: US 60/113,439
; PRIOR FILING DATE: 1998-12-01

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 936
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; US-09-452-380-3

Query Match 41.2%; Score 1965; DB 9; Length 936;

Best Local Similarity 43.4%; Pred. No. 1.6e-138;
Matches 411; Conservative 179; Mismatches 327; Indels 30; Gaps 16;

Qy 1 MKSSPPKVFSTFALP-PLSMIATETVLDSS-ASPDGKNGNFSVRESQE-DAGTYYLFK 57
Db 1 MKSSVSWLFFSIIPLFSSLSIVAAEVTLDDSSNNSYDGSNGTTFVFTTDAAGTYYLL 60
Qy 58 GNVTLLENIPGTGTATTKCFNNTKGDLTFTGNGNSLLFQTVDAGTVACAAVNSSVVDKST 117
Db 61 SDVSTQACALGIPLASGCFLEAGDLTFQGNQHALKAFAPINAGSSAGTVAASADKXL 120
Qy 118 TFIGFSSIFIASPGSSIT-TGKGAVSCSTGSLSTKVNLSLFFSKNFSTNDGGAITAKTL 176
Db 121 LFNDFSLIISCPILLSPTGCCALK-SVGNLSLTGNSQIIFTQNFSSDNGGVINTKNF 179
Qy 177 SLTGTMSALFSENT--SSKKGAIQTSALTITGNQGEVSFSDNTSSDGAALFTEASV 234
Db 180 LUSGTSQAFSFRNQAFQTKQGVVYATGTITIENSPIGVFSQNLAKSGGALYSTDNC 239
Qy 235 TISNNAKVFIQNKVTGASSSTTGDMSGGAICAYKTSDDTKVTLTGQNMLLFSNNTSTA 294
Db 240 SITDNFQVIFDGNASWEAQA-----QGAICC--TTTDKTVTLTGKNSLFTNNALTY 292
Qy 295 GGAIVYKLEASGLTLFERNVNGTAPK--GGAIAIEDSGELSLSDSDGDIIVFLGN 352
Db 293 GGAISGLKVSISAGGPTLFSN-1SGSAGQGGGAINIASAGELALSATSGDITFNQ 351
Qy 353 VSTTPTGNNRSIDLTGSAKMTALRAAGRAIYFYDPIITGSSSTTVDLKYNETPADS 412
Db 352 VINGSTST-RNAINIIDTAKTISIRATGQSIYFYDPIITGTAASTDTLNLNADANSE 410
Qy 413 LQYTGNIITFGKLSSETAADSNNLSKLLQPVTLSSGGLSLKHGVTLOTQFTQOADS 472
Db 411 IEYGAIVPSGKLSPEKAIAANVTSTIRQPAVLARGDLVLRGCVTVTFKDLTQPSGR 470
Qy 473 LEMDVGTTLPEADTS--TINNIVINISIDGAKAKIETKATSKNLTLSGTTITLDDPTGTF 531
Db 471 ILMGDTTLISAKENLNLGLAVNLSSLDGTNKAALKEADKNISLSTGTLALIDTSGSF 530
Qy 532 YENHSLRNPQSYDILELK---ASGTVTSTAVTPDPIMGEKPHYGYQGTGPIVWGTGAST 588
Db 531 YENHSLKASATYPLELTTAGANGITTLGALSTLALQBPETHYGYQGNW-QLSWANATSS 589
Qy 589 -TATFNWTKTYIPAPERIGSLVPSNLWNAFIDISSLHYLMETANEGLQDRAFWCAGLS 647
Db 590 KIGSINWTRTGYTPSPERKSNLPLNSLMGNFDIRSINQLIBTKSSGPPFERELWSGIA 649
Qy 648 NPFHKDSTKTRRGFRHLSCGYVIGNLTSCDKILSAFQCFGRDRDYFVAKNQGTGTYG 707
Db 650 NPFYRDSMPTRGFRHISGGYALGLTATTPADQUTFAFCQLFADRNHIITGKNGHDYIG 709
Qy 708 GTLYYQHNETHYISLPCKL-----RPSLSVYVPTTEIIVLFSNLSYTHNDNLKTKYTY 761
Db 710 ASLYFHTHTEGLFDIANFLMGKATRAPVLSLSEISQIPLSFOAKFSYLTNDHMKTYTND 769
Qy 762 PTVKGSGWNSDPALEFGGRAPICLDESALFEQYMFPMKLOFVVAHQGFKEQGTAREFG 821
Db 770 STIKSGWRNDAPCADLGASLPFVIVPVLKKEVEFVKQVYIAHQQDFYERHAEGRAF 829
Qy 822 SGRVNLALPIGRFRDKESDCQDATYNLTGYTVDLVRSNPDCTTLIRISGDSWKTFTG 881
Db 830 KSELINVEIPIGVTTERDSKSEKGVLDLTMLYILDAYRNPCKQTSLLASDANWYATGN 889
Qy 882 LARQALVLRAGNHFCFNSNFEAFSPFELRGSSRNVDLGAKYQF 928


```

Db 300 PQSN--ISGSSAGQGGGAINIASAGELALSATSGDITFNNQVNGSTST-RNAINIIDT 357
Qy 371 AKMTALRAAGRAIYFYDPIITGSGTTVDLVKNETPADSALQVGTGNIIFTEGKLSETE 430
Db 358 AKVTSIRAATGOSIYFYDPIITGSGTTVDLVKNETPADSALQVGTGNIIFTEGKLSETE 417
Qy 431 AADSKNLSKLLQPVTLSSGGLTSLKHGVTLOQAFQOQADSRLEMDVGTTLLEPADTS-TI 489
Db 418 KAIAANVTSTIROPAVLARGDLVLDGVTVPFKOLTQSPGRILMDGGTTLISAKENLSL 477
Qy 490 NNLVINISIDGAKAKIETKATSKNLTSLGTTILDPGTGYENHSLRNPOSYDILELK 549
Db 478 NGLAVNLSLDTGNTKAALKEAADKNISLGSITLIDTSGSYENHSLKASTYPLELLELT 537
Qy 550 ---ASGTVTSTAVTDPDPIINGEKFHYQGTWGPVIMVGTGAST-TATFNWTKTYIPNPER 605
Db 538 TAGANGTITLGAISTLTLOQEPETHYQGNW-QLSWANATSSKIGSINWTRTGYIPSPER 596
Qy 606 IGSVLPNSLWNAFIDISSLHYLMETANEGLQDRAFWCAGLSNFFHKSOTKTRRGRHLS 665
Db 597 KENLPLNSLWGNFIDIRSINQLIETKSGEPFERELMWSGIANFFYRDSMPTRHGFRHS 656
Qy 666 GGYVIGGNLHTCSDKILSAFQOLGDRDRDYFAKNQGTGYGGLTYQHNETHYISLPCKL 725
Db 657 GGYALGITATTAPEDQLTFAFQOLFARDNRNHTGKNHGDYTGASLYFHHTEGLFIANFL 716
Qy 726 -----RPSLSVVPTEIPVLSGNLSYTHDNDLTKYTYTTPVKSGMNDSPFALFEGG 779
Db 717 WKGATRAPVWLSEISOIIPLSFPAKFSYLHTDNHMKTYTIDNSIIGKSWRNDFAFCADLGA 776
Qy 780 RAPICLDESALFQYMPFVKQVIAHQDQFYERHAEGRFNAKSELINVEIPIGVTFERD 836
Db 777 SLFPVISVPLLKEVEFPVKQVIAHQDQFYERHAEGRFNAKSELINVEIPIGVTFERD 836
Qy 840 SDCQDATYMLTGYTVDLVRSDNCDCTTLRISGDSWKTGTGNLAROALVLRAGNHFCFNS 899
Db 837 SKSEKGYDITLMIYILDVARNPKQCTSLIASDANWMAYGATNLARQFVSRAANHFQVNP 896
Qy 900 NFEAFQSFPELGRSSRNYNVDLGAQYF 928
Db 897 HMEIFQFAFEVRSRNNYNTNLGSKPCF 925

```

```

RESULT 5
US-10-324-129-4
; Sequence 4, Application US/10324129
; Publication No. US20030157124A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN ET AL.
; TITLE OF INVENTION: Chlamydia antigens and corresponding DNA fragments and uses thereof
; FILE REFERENCE: 77813-4
; CURRENT APPLICATION NUMBER: US/10/324,129
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/113,439
; PRIOR FILING DATE: 1998-12-01
; PRIOR APPLICATION NUMBER: US 60/132,272
; PRIOR FILING DATE: 1999-05-03
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 925
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-324-129-4

```

```

Query Match 40.8%; Score 1946.5; DB 12; Length 925;
Best Local Similarity 43.5%; Pred. No. 3.8e-137;
Matches 404; Conservative 177; Mismatches 319; Indels 29; Gaps 15;
Qy 18 LSMATETVLDSS-ASPDGKNGNFSVRESOE-DAGTTLFKGNVTLENIPCTGTATKS 75
Db 8 LSIIVAAEVTLDSSNNNSYDGSNGTTFVTFTDAAAGTTYSLLSDVSFQAGALGIPLASG 67

```

```

Qy 76 CFNNTKGLDTFTGNSLLFQTVDAAGTVAGAAVNSSVVDKSTTTFIGFSSLSFIASPGSSI 135
Db 68 CFLEAGDGLTFQGNQHALKAFINAGSAGTVASTAADKNLLENDFSRSLISICPSULL 127
Qy 136 T-TCKGAVSCSTGSLSTLKNVSLLSFKNSFTDNGGAIKTAKTSLTGTMTGALFSENT--S 192
Db 128 SPTQCCALK-SVGNLSLTGNSQIIFTFNFSDDNGGVINTKNFLSLGTSQSPASFRNQAPT 186
Qy 193 SKKGGAQTSDALITGNQGEVSDNTSSDGAITEASVTISNNAKVSFIDNKVTGA 252
Db 187 GKQGGVYATATTIENSPIGVSPSONLAKSGGALYSTDNCSITDNPQVIFDGNASWEA 246
Qy 253 SSSITGDMGSGCAICAYKTSITDKVTLTGCMOLLSFNNTTGTAGGAIYVKKLELASGGLTL 312
Db 247 AQA-----QGAICG--TTTDTKTTLGNKNLSFNTNLTALTYGGAISGLKVISIAGGPTL 299
Qy 313 FSRNSVNGGTAPK--GGAIAIEDSGELSDSADSGDIVELGNVTSTTGTNRSSIDIGTS 370
Db 300 PQSN--ISGSSAGQGGGAINIASAGELALSATSGDITFNNQVNGSTST-RNAINIIDT 357
Qy 371 AKMTALRAAGRAIYFYDPIITGSGTTVDLVKNETPADSALQVGTGNIIFTEGKLSETE 430
Db 358 AKVTSIRAATGOSIYFYDPIITGSGTTVDLVKNETPADSALQVGTGNIIFTEGKLSETE 417
Qy 431 AADSKNLSKLLQPVTLSSGGLTSLKHGVTLOQAFQOQADSRLEMDVGTTLLEPADTS-TI 489
Db 418 KAIAANVTSTIROPAVLARGDLVLDGVTVPFKOLTQSPGRILMDGGTTLISAKENLSL 477
Qy 490 NNLVINISIDGAKAKIETKATSKNLTSLGTTILDPGTGYENHSLRNPOSYDILELK 549
Db 478 NGLAVNLSLDTGNTKAALKEAADKNISLGSITLIDTSGSYENHSLKASTYPLELLELT 537
Qy 550 ---ASGTVTSTAVTDPDPIINGEKFHYQGTWGPVIMVGTGAST-TATFNWTKTYIPNPER 605
Db 538 TAGANGTITLGAISTLTLOQEPETHYQGNW-QLSWANATSSKIGSINWTRTGYIPSPER 596
Qy 606 IGSVLPNSLWNAFIDISSLHYLMETANEGLQDRAFWCAGLSNFFHKSOTKTRRGRHLS 665
Db 597 KENLPLNSLWGNFIDIRSINQLIETKSGEPFERELMWSGIANFFYRDSMPTRHGFRHS 656
Qy 666 GGYVIGGNLHTCSDKILSAFQOLGDRDRDYFAKNQGTGYGGLTYQHNETHYISLPCKL 725
Db 657 GGYALGITATTAPEDQLTFAFQOLFARDNRNHTGKNHGDYTGASLYFHHTEGLFIANFL 716
Qy 726 -----RPSLSVVPTEIPVLSGNLSYTHDNDLTKYTYTTPVKSGMNDSPFALFEGG 779
Db 717 WKGATRAPVWLSEISOIIPLSFPAKFSYLHTDNHMKTYTIDNSIIGKSWRNDFAFCADLGA 776
Qy 780 RAPICLDESALFQYMPFVKQVIAHQDQFYERHAEGRFNAKSELINVEIPIGVTFERD 839
Db 777 SLFPVISVPLLKEVEFPVKQVIAHQDQFYERHAEGRFNAKSELINVEIPIGVTFERD 836
Qy 840 SDCQDATYMLTGYTVDLVRSDNCDCTTLRISGDSWKTGTGNLAROALVLRAGNHFCFNS 899
Db 837 SKSEKGYDITLMIYILDVARNPKQCTSLIASDANWMAYGATNLARQFVSRAANHFQVNP 896
Qy 900 NFEAFQSFPELGRSSRNYNVDLGAQYF 928
Db 897 HMEIFQFAFEVRSRNNYNTNLGSKPCF 925

```

```

RESULT 6
US-09-738-269-57
; Sequence 57, Application US/09738269
; Publication No. US20030185848A1
; GENERAL INFORMATION:
; APPLICANT: JOHNSTON, STEPHEN A.
; APPLICANT: STEWKE-HALE, KATHERINE
; APPLICANT: SKES, KATHRYN P.
; APPLICANT: KALTENBOECK, BERNHARD
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
; NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF

```

```
; TITLE OF INVENTION: CHLAMYDIA PSITTACI
; FILE REFERENCE: UTSD:659
; CURRENT APPLICATION NUMBER: US/09/738,269
; CURRENT FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
; LENGTH: 926
; TYPE: PRT
; ORGANISM: Chlamydia psittaci
US-09-738-269-57

Query Match      40.1%; Score 1915; DB 12; Length 926;
Best Local Similarity 42.9%; Pred. No. 8.7e-135; Indels 56; Gaps 18;
Matches 410; Conservative 156; Mismatches 333;

QY 1 MKSFPKVFSTFAIFPLSM-----IATETVLDSSASFDGKNGNFSVRESQEDA 50
DB 1 MRSLYKILISSLTLPISFHSQHAELVLTQESILDA-----NGAFSPQSTSTAG 52

QY 51 GTTYLFGKGNVTLENIPTGTATATKSCFNNTKGLDFTFGNGNSLLFQTVDAAGTAVAGAAVNS 110
DB 53 GTTYNVESDISIVDV-QQTAALASSAFVQADNLTFKGNHSLSIITNANAG-ANPAGINV 110

QY 111 SVVDKSTTFIFGFSLSFIAPSGSIITGKAVSCSTGSLTKNVSLLFSKNFSTDNCGA 170
DB 111 NTADKILTLTDFSKLSFKECPSSLVNTGKGMK-SGGALNLANNASILFQNTYSAENNGA 169

QY 171 ITAKTLTLTGTMTSALFSENTSSKKGAIQTSALITITGNQGEVSFSDNTSSDGAIAFT 230
DB 170 ISCKAFSLTGSSEKISFTTNTAKKGAIAATGIAHLSDNQGTIRFSGNTAVNSGGAVYS 229

QY 231 EASVTISNNAKVFIIDNKVITGASSTTGDMSGGAI CAYKTSITDKVTLTGQMLLFNSNT 290
DB 230 EASMTIAGNNHVAFNNNAVSGSS-----DGGCGAIHCSKTGSAPTLTIRDNKVLIFEENT 284

QY 291 STTAGGAIYVKKLELASSGGLTLFSRNSVNGTAPKGAIAIEDSGELSLSDSDGDIIVFLG 350
DB 285 SSARKGGAITYDKLTLTSGGPTAFINNKVTHAT-PRKGAIGIAANGECSLTAHEDGDIITFDN 343

QY 351 NTVTSTTPGT-NRSSIDLGTSAKMTALRSAGRAIFYDPDITGSSSTTVTDVLKVNTPA 409
DB 344 NLMATQDNATIKRNAINEGKFNVLRAASGKTI SFYDPIITVEGN--AADLLTLNKAEG 401

QY 410 DSALQYTGNIIFTGEKLUSETAADSKNLTSKLPQVTLTSGGTLSLKHGVTLOQAFVQQA 469
DB 402 DKT--YNGRIIFSGEKLTEEQAAVADNLKTTFTQPIITLAAGELVLRSGVEAEKTVVQTA 459

QY 470 DSRLEMDVGTLE-PADTSTINNLVINISSIDGAKKAKIETKATSKNLTLSGTLTLDPT 528
DB 460 GSLILMDAGTKLSAKTADATLNLAINPNTLDGKFPVAVDAVAKNVTLSGALGVIDPT 519

QY 529 GTFVENHSLRNPSQSDILELKASGTVTSTAVTPDPIMG-EKPHYGYQGTWGPVW----- 582
DB 520 GKPHYENHKLNDLALGGIQLSGKSVITTVN-PSHVVGVAETHYGYQGNWS-VSWVKDNN 577

QY 583 GTGASTATFNWTKTGYIPNPERIGSLVPNSLWAFIDISLHLVIMETANEG-IQGBRAF 641
DB 578 SDPKTQTAIFTWNTKTYVNPERRAPLVNLSWGSFIDLRSIQDVLERSVDLSILETRGL 637

QY 642 WCAGLSNPFPHKDSKTRGFRHLSGGYVIGGNLHCTCSKILSAAPCOLFGDRDRDYFAKN 701
DB 638 WSGIGNEFFHKDRNAENKFRHISGVIIGATTNFRSDLSVAFQCFADKDKYLVSKN 697

QY 702 QGTVGGTLYYQHNETHYISL-----PCKLRPCSLSYVPTPEIPVLFSGNLSYTHTDND 753
DB 698 AANYVAGSVYQHVSKFDDLTFLNPNPTC-----CSGFSKEIPIFLDAQIITYCHTANN 751

QY 754 LKTYTTPYTVKSGWNSDFALEFGGRAPICLDSALPEQWPMFKLOFVIAHQEGFKBQ 813
DB 752 MTTSTYDYPEVKSGWNTGLTITLSTSVPIFVFSSIFDSYAPFAKLQVYAHQDDFKEP 811

QY 814 GTEAREFGSSRLVNLALPIGIRFDKESDCQDATYNLTIGYTVDLVRSNPDCTTTILRISGD 873
```

```
DB 812 TTTEGRVFESSDLLNVSPIGIKFEKLSYCGERSAYDLTLMYIPDVYVRHNPSCMTGLAINDV 871
QY 874 SWKTGNTLNARQALVLRAGNHCFNSNFEAFQSFELRGSSRNYNVDLGAQVQF 928
DB 872 SWLTATNLARQAFIVRAGNHIALTSGVEMFQSFQFELRSSRNYNVDLGAQVAF 926

RESULT 7
US-10-023-437-57
; Sequence 57, Application US/10023437
; Publication No. US20020183272A1
; GENERAL INFORMATION:
; APPLICANT: JOHNSTON, STEPHEN A.
; APPLICANT: STEWKE-HALE, KATHERINE
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: KALTENBOECK, BERNHARD
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING NUCLEIC ACID
; TITLE OF INVENTION: AND/OR POLYPEPTIDE SEQUENCES OF CHLAMYDIA
; FILE REFERENCE: UTSD:736US
; CURRENT APPLICATION NUMBER: US/10/023,437
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/225,839
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
; LENGTH: 926
; TYPE: PRT
; ORGANISM: Chlamydia psittaci
US-10-023-437-57

Query Match      40.1%; Score 1915; DB 14; Length 926;
Best Local Similarity 42.9%; Pred. No. 8.7e-135; Indels 56; Gaps 18;
Matches 410; Conservative 156; Mismatches 333;

QY 1 MKSFPKVFSTFAIFPLSM-----IATETVLDSSASFDGKNGNFSVRESQEDA 50
DB 1 MRSLYKILISSLTLPISFHSQHAELVLTQESILDA-----NGAFSPQSTSTAG 52

QY 51 GTTYLFGKGNVTLENIPTGTATATKSCFNNTKGLDFTFGNGNSLLFQTVDAAGTAVAGAAVNS 110
DB 53 GTTYNVESDISIVDV-QQTAALASSAFVQADNLTFKGNHSLSIITNANAG-ANPAGINV 110

QY 111 SVVDKSTTFIFGFSLSFIAPSGSIITGKAVSCSTGSLTKNVSLLFSKNFSTDNCGA 170
DB 111 NTADKILTLTDFSKLSFKECPSSLVNTGKGMK-SGGALNLANNASILFQNTYSAENNGA 169

QY 171 ITAKTLTLTGTMTSALFSENTSSKKGAIQTSALITITGNQGEVSFSDNTSSDGAIAFT 230
DB 170 ISCKAFSLTGSSEKISFTTNTAKKGAIAATGIAHLSDNQGTIRFSGNTAVNSGGAVYS 229

QY 231 EASVTISNNAKVFIIDNKVITGASSTTGDMSGGAI CAYKTSITDKVTLTGQMLLFNSNT 290
DB 230 EASMTIAGNNHVAFNNNAVSGSS-----DGGCGAIHCSKTGSAPTLTIRDNKVLIFEENT 284

QY 291 STTAGGAIYVKKLELASSGGLTLFSRNSVNGTAPKGAIAIEDSGELSLSDSDGDIIVFLG 350
DB 285 SSARKGGAITYDKLTLTSGGPTAFINNKVTHAT-PRKGAIGIAANGECSLTAHEDGDIITFDN 343

QY 351 NTVTSTTPGT-NRSSIDLGTSAKMTALRSAGRAIFYDPDITGSSSTTVTDVLKVNTPA 409
DB 344 NLMATQDNATIKRNAINEGKFNVLRAASGKTI SFYDPIITVEGN--AADLLTLNKAEG 401

QY 410 DSALQYTGNIIFTGEKLUSETAADSKNLTSKLPQVTLTSGGTLSLKHGVTLOQAFVQQA 469
DB 402 DKT--YNGRIIFSGEKLTEEQAAVADNLKTTFTQPIITLAAGELVLRSGVEAEKTVVQTA 459

QY 470 DSRLEMDVGTLE-PADTSTINNLVINISSIDGAKKAKIETKATSKNLTLSGTLTLDPT 528
DB 460 GSLILMDAGTKLSAKTADATLNLAINPNTLDGKFPVAVDAVAKNVTLSGALGVIDPT 519

QY 529 GTFVENHSLRNPSQSDILELKASGTVTSTAVTPDPIMG-EKPHYGYQGTWGPVW----- 582
```

```

Db      520 GKFEYENKLANDTLALGQIOLSGKGSVTTNNV-PSHVVGVVAETHYGYQGNWS-VSWVKDNN 577
Qy      583 GTGASTTATENWTKGTGVPINPERIGSLVPNSLWNAFIDISLHLYMETANEG-LQGDRAF 641
Db      578 SDPKTQTAIFWNTKGTGVPINPERAPLVLNLSGFSFIDLSRIQDVLERSVDLSILETRGL 637
Qy      642 WCAGLSNFFHKDSTKTRGPHRLSGGVYVGNLHTCSDKILSAAPCOLFGDRDRYFVAKN 701
Db      638 WUSGIGNFFHKDRNAENRKPRISSGVYLGATTNSTRSDLSVAFCOLFAKDKDYLVSKN 697
Qy      702 QGTVYGGTLYYQHNETYISL-----PCKLRPCSLSYVTEIPVLPSGNLSYTHTDND 753
Db      698 AANYAGSVYVYQHSKFDLTLFNGPNTC-----CSGFSKEIPILDAQITCYHTANN 751
Qy      754 LKTYTTPYTVKSGWGNDSFALEFGGRAPICLDESALFEQWMPFMKQFVYAHQEGFKEQ 813
Db      752 MTSYTDYPEVKGSGWGNDTLGLTSLTVPIPVFSSSIFDSYAPAKLVVYAHQDDFKEP 811
Qy      814 GTEAREFGSSRLVNLALPIGIRFDEKESQCDATYNLTLGYTVDLVRSNPDCTTLIRISGD 873
Db      812 TTEGRVPFSSDLLNVSVBIGIKFEKLSYGERSAYDLTLNYPDVYRHNPSQMTGLAINDV 871
Qy      874 SWKTFGTNLARQALVLRAGNHCFSNFEAFSOFSEFELRGSSRNYNVDLGAKYQF 928
Db      872 SWLTATNLARQAFIVRAGNHIALTSGVEMFSGFELRGSSRNYNVDLGAKVAF 926

RESULT 8
US-09-738-269-23
; Sequence 23, Application US/09738269
; Publication No. US20030185848A1
; GENERAL INFORMATION:
; APPLICANT: JOHNSTON, STEPHEN A.
; APPLICANT: STEMKE-HALE, KATHERINE
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: KALTENBOECK, BERNHARD
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
; NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF
; CHLAMYDIA PSITTACI
; FILE REFERENCE: UTSD:659
; CURRENT APPLICATION NUMBER: US/09/738,269
; CURRENT FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 23
; LENGTH: 839
; TYPE: PRT
; ORGANISM: Chlamydia psittaci
US-09-738-269-23

Query Match      34.7%; Score 1656.5; DB 12; Length 839;
Best Local Similarity 40.3%; Pred. No. 1.8e-115;
Matches 381; Conservative 133; Mismatches 285; Indels 147; Gaps 24;

Qy      11 STFAIPPLSMI-ATETVLSSASFDGN-KNGNSVRESQEDACTTFLPKGNVLENIPTG 68
Db      13 SLFASNSLSPANDAQTALTPSDSYNGNVTSEFPQVRET--SSGTTYTCGNVCI-SFAGK 69
Qy      69 GTAITKSCFNNTKGLTFTGNGSNLLFQTVDACTVAGAAVNSVVDKSTTFFGSSLSFI 128
Db      70 DSGLKKSFCFAT-DNLTFLNGVYTLCPDNITTTASNPAGINVOGQKTLGISGSLFSCA 128
Qy      129 ASPGSSITTKGAVSCSTGSLSTKLVSLFKNSFSTDNGGAIKTAKTSLTGTMTSALFS 188
Db      129 YCPGP--TTGVALQ-TKGNITLKDNSSLVFKHNCSTAEAGAIQC----- 170
Qy      189 ENTSSKKGAIQTSALITITGNQGEVSFSDNTSSDGAALFTEASVTISNNAKVSIIDNK 248
Db      171 ----- 170
Qy      249 VTGASSSTTGDMSSGALCAVKTSTDTKVTLTGNQMLFNSNTSTTAGGAIYVKKLELASG 308

```

```

Db      171 -----KGSSDAELKIENNQNLFVSENSSTSKGAIYADKLTIIVSG 210
Qy      309 GLTLFRNSVNGTAPKGAIAIED-SGELSLSADSGDIVFLGNVTSTTPGTN---RSS 364
Db      211 GPTLFNNSVNSGSSPKGGAISIKDSGSECSLADLGDITFDGNKIKTSGGSSTVTRNS 270
Qy      365 IDLGTSAKMTALSAAGRAIFYDPITTTGSSSTTVTVLVKNVETPADSALQYTGNIIFTGE 424
Db      271 IDLGT-GKFTKLAKDGFIFFFDPIITGGG-----DELNINK---KETVDYTKIVFSGE 322
Qy      425 KLSETEADSKNLTSLKLPVLQVLTSGGTLKLGHTVLTQTAFTQQAOSRLMDVGTTLLEP- 483
Db      323 KLSDEEKARAENLASTFNQPIITLSAGSLVLKDGVSFTAKQVTEAGSTVVMDLGTLTQTP 382
Qy      484 ---ADISTINNVLINISSI---DGAKKAKIETKATSKNLTSLGTTITLLDPTGFYENHSL 537
Db      383 SSGGETITTLNLDINTASLGGGGTSPAKLATNTASQAITIN-AVNLDVADGNAIEDPIL 441
Qy      538 RNPOSYDILBLKASGTVTSTAIVDPDPIMG- --KFHYGYQGTWGPVIMVGTGAST-TATEN 593
Db      442 ATSKPFTAI---VATTNASTVTVQTDNLTVNVPPTHYGYQGNM-TVTWDTETATKATLT 497
Qy      594 WTKTGYIPNPERIGSLVPNSLWNAFIDISLHLYMETANEGLOCDRAFWCAGLSNFFHKD 653
Db      498 WEQTGYSPNPERQGPLVNTLWGAFLSALRAIQNLMDISVNGADYHRCGFVWSGLANFLHKS 557
Qy      654 STKTRGFRHLGGYVTLGGNLHTCSDKILSAAPCOLFGDRDRYFVAKNQGTVYGGTLYYQ 713
Db      558 GSDTKRFRHNSAGYALGVYAKTSPDDIFSAAPCOLFGKDKDYLVSNNANIYAGSLYYQ 617
Qy      714 HNETYISLPCKLRPCSLSYVTEIPVLPSGNLSYTHTDNLT-----KYTTYPTVK 765
Db      618 H-----ISWSAWQNLLQNTIGAEAPLVNLAQLTYCHASNDMKTNTTTPYAPKTYAEIK 673
Qy      766 GSWGNDSPALEFGGRAPICLDESALFEQWMPFMKQFVYAHQEGFKEQGT-EAREFGSSR 824
Db      674 GDWGNDCFGVELGATVPIQTSSLLFDMYSPFLKFLVHTHQDDDFKENNSDQGRYFESSN 733
Qy      825 LVNLALPIGIRFDEKESQCDATYNLTLGYTVDLVRSNPDCTTLIRISGDS--WKTFTGNL 882
Db      734 LTNLSLPIGIRFERPANNDTASYHVTAAYSPDIVRSNPDCTTSLLVSPDSAVVWTKANL 793
Qy      883 ARQALVLRAGNHCFSNFEAFSOFSEFELRGSSRNYNVDLGAKYQF 928
Db      794 ARSAFMLQAGNVLSHNIEIFSGFELRGSSRNYNVDLGSKIOF 839

RESULT 9
US-10-023-437-23
; Sequence 23, Application US/10023437
; Publication No. US20020183272A1
; GENERAL INFORMATION:
; APPLICANT: JOHNSTON, STEPHEN A.
; APPLICANT: STEMKE-HALE, KATHERINE
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: KALTENBOECK, BERNHARD
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING NUCLEIC AC
; AND/OR POLYPEPTIDE SEQUENCES OF CHLAMYDIA
; FILE REFERENCE: UTSD:736US
; CURRENT APPLICATION NUMBER: US/10/023,437
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/225,839
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 23
; LENGTH: 839
; TYPE: PRT
; ORGANISM: Chlamydia psittaci
US-10-023-437-23

Query Match      34.7%; Score 1656.5; DB 14; Length 839;
Best Local Similarity 40.3%; Pred. No. 1.8e-115;

```



```
Matches 381; Conservative 133; Mismatches 285; Indels 147; Gaps 24;
QY 11 STEAIFPLSMI--ATETVLDSSASPDGN--KGNFNSVRESQEDAGTTLFKGNVTLENIGPT 68
Db 13 SLFASNSLSFANDAQTALTTPDSYNGNVTSEFQVKE--SSGTTYTCEGNVCI--SPAGK 69
QY 69 GTAITKSCFNNTKGDLTFTGNGNSLLPQTVDAGTVAGAAVNSVVDKSTFTFGFSGSLFI 128
Db 70 DSGLKSCFSAT--DNLTFNGNYTLCFDDNITTAENPGAINVQOGQKTLGIGSFLFSCA 128
QY 129 ASPGSSITTKGAVSCSTGSLSTKNVSLLFKNFSTDNNGGAIKTATLSLTGTTMSALFS 188
Db 129 YCPGP--TTGYGAIQ--TKGNTTLKONSSLVFHKNCSTABGGAIQ-- 170
QY 189 ENTSSKKGAIQTSDALTITIGNQGEVFSFSDNTSSDGAALFTEASVTISNNAKVFDINK 248
Db 171 ----- 170
QY 249 VTGASSITGDMGGAI CAYKSTDTKVTLTGQMLLFSNNTSTTAGGAIYVKKLELASG 308
Db 171 -----KGSDAEKLENNQNLVFSNSSTSKGGAAYADKLTVSG 210
QY 309 GLTLFNSVNGTAPKGAIAIED--SGELSLSADSGDIVFLGNTVTSTPGTN---RES 364
Db 211 GPTLFNSNSVNGSPKGAISIKDSGRCSLTADLGITFDGKNIKTSGGSTVTRNS 270
QY 365 IDLGTSAKMTALRSAAAGAIYFVDPTTGGSTTVTVLKVNETPADSALOYTCNLIPTGE 424
Db 271 IDLGT--GKTKLRADGFGFFDYDPTTGGGS---DELINK---KETVDYTKGVFSGE 322
QY 425 KLSSETAADSKNLTSKLLQPVTLSCGTLASIKHGVTLQTOAFTQOADSRLMDVGTLEP- 483
Db 323 KLSDEEKAENLSTFNPOTILSAGSLVKGVSTAKQVTOEAGSTVVMDLGTLTLP 382
QY 484 ---ADTSTNNLVINISSI---DGAKAKIETKATSKNLTSLGCTITLLDPTGTFYENHSL 537
Db 383 SSGEITITLNDINIASLGGGGTSPAKLATNTASQAITIN--AVNLVDADGNAYEDPIL 441
QY 538 RNPQSDIILELKASGTVTSTAVTPDPMGE---KFHYGQGTWGPVWGTGAST--TATFN 593
Db 442 ATSKPFTAI---VATTNASTVTQTDNLNTVVPPTHYGYQGNW--TVTMDTETATKTATLT 497
QY 594 WTKGTIPNPERIGSLVPSNLWAFIDISLHLYMETANEGLQDRAFCAGLSNPFHKD 653
Db 498 WEQTGSPNERQGLVNTLWAGFSDLRATQIONLMDISLVNGADYHGRGFWVSGLANFLHKS 557
QY 654 STKTRRGFRHLSGGYVIGGNLHCTSKILSAAFCQLFGRDRDYFVAKNQGTVYGGTLYYQ 713
Db 558 GSUTKKFRNSAGYALGVYAKTPSDDI FSAAFQCLFGKDKDYLVSNNANIYAGSLYYQ 617
QY 714 HNBTYISLPCKLAPCSLSYVPTTELPVLPSGNLSYTHTDNDLKT-----KYTTYPTVK 765
Db 618 H---ISYWSAWQMLLQNTICAEAPLVNLAQLTYCHASNDMKNTMTTYAPRKTYYABEK 673
QY 766 GSWGNSFALEFGGRAPICLDESALPEQYMPFKLPVYAHQGEFKQGT--BAREFGSSR 824
Db 674 GDMWDCFGVELGATVPQIQTSSLLFDMYSEFLKQLVHTHQDDFKENNSDQGRYFESSN 733
QY 825 LVNLALPIGRFKESDCODATVNLTLGYTVDLVRSNPDCTTTHIRISGDS--WKTFTNL 882
Db 734 LTNLSLPIGIKFERFANNDTASYHTVTAAYSPDIVRSNPDCTTSLVSPDSAVVWTKANL 793
QY 883 ARQALVLRAGNHCFNENFAFOSFELRGSSRYNVVDLGAKYQF 928
Db 794 ARSAFMLQAGNYSLSLHNIIRIFSGFELRGSSRTYVNDLGSKIQF 839
```

RESULT 10

US-09-886-468-19
; Sequence 19, Application US/09886468
; Patent No. US20020037293A1
; GENERAL INFORMATION:
; APPLICANT: Aventis Pasteur Limited

```
; TITLE OF INVENTION: Chlamydia antigens and corresponding DNA fragments and uses the;  
; FILE REFERENCE: 77813-5  
; CURRENT APPLICATION NUMBER: US/09/886,468  
; CURRENT FILING DATE: 1999-12-23  
; PRIOR APPLICATION NUMBER: 60/113,280  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/113,281  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/113,282  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/113,283  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/113,284  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/113,285  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/113,385  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/114,050  
; PRIOR FILING DATE: 1998-12-28  
; PRIOR APPLICATION NUMBER: 60/114,056  
; PRIOR FILING DATE: 1998-12-28  
; PRIOR APPLICATION NUMBER: 60/114,057  
; PRIOR FILING DATE: 1998-12-28  
; PRIOR APPLICATION NUMBER: 60/114,058  
; PRIOR FILING DATE: 1998-12-28  
; PRIOR APPLICATION NUMBER: 60/114,059  
; PRIOR FILING DATE: 1998-12-28  
; PRIOR APPLICATION NUMBER: 60/114,061  
; PRIOR FILING DATE: 1998-12-28  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 19  
; LENGTH: 922  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
US-09-886-468-19
```

Query Match 30.1%; Score 1437.5; DB 9; Length 922;
Best Local Similarity 36.8%; Pred. No. 5.6e-99;
Matches 350; Conservative 156; Mismatches 382; Indels 63; Gaps 26;

```
QY 5 FP-KFVFSTFAIFPLPSMIATETVLDSSASFDGKNGNFSVRESQEDAGTTLFKGNVTLE 63  
Db 8 FPLVFSFTLLSVFDTLSLATTISLTPEDSFHGDSQN--AERSVNVQAGDVYSLTGDVSL 65  
QY 64 NIPGTGTAITKSCFNNTKGDLTFTGNGNSLLPQTVDAGTVAGAAVNSVVDKSTT-FIG 122  
Db 66 NV--DNSALNKAFCFNVTSVTSVTFAGNHGGLYFNNTISSGTTKEGAVLCCQDPQATAPSGF 123  
QY 123 SLSLPTASPGSSITTTCKGAVSC--STGSLSLTQVSLFQKQFSTDNCGAITAKTILSLTG 180  
Db 124 STLSFQSPGDIKEQ-----CLYSKVALMLNLYVVRFEQNSKTKGGAISGANTVIG 178  
QY 181 TTMSALFSENTSKKGAIQTSDALTITIGNQGEVFSFSDNTSSD--SGAALFTEASVTISNN 239  
Db 179 NYDSVSVFYQMAAT--FGAITHSSGPIQIAVNOABIREAQTAKNGSGALYSDGDIIDQN 237  
QY 240 AKVSPFDNKVTVGASSTTGMDSGGAICAYKTSTDTK-----VTLTGQMLLFSNNTSTAG 295  
Db 238 AYLVLFRENE-----ALTTAIGKGGAVCCLPSTGSGSTPVPITVTFSDNKQLVFEENHIMG 292  
QY 296 GAIYVKKLELASGLTLFERNVNGTAPKGAIAIEDSGELSLSADSGDIVFLGNTVTS 355  
Db 293 GAIYARKLSISSGGPTLFINNISYANSONLGAIAIDTGEISLSBAKGTIITQGN--RT 350  
QY 356 TTPGTNRSSIDLGTSAKMTALRSAAAGAIYFVDPTTGGSTTVTVLKVNETPADSALOY 415  
Db 351 SLPLFN--GHLLQNAKFLKLAQNGVSYIEFYDPI--SEADGSTQLNINGDPKKN--EY 404  
QY 416 TGNIIFTGEKLSSETAADSKNLTSKLLQPVTLSCGTLASIKHGVTLQTOAFTQOADSRLM 475  
Db 405 TGTILFSGEK---SLANDPRDFKSTIPQNVNLSAGVILVIKEGAETVTSKFTQSPGSHVL 461
```


Db 580 VANTLWNTYSDMQAVQSMINTTAHG--GAYLFTGWSAVSNLFYVHDSGSKPIDNWHRS 637
Qy 666 GGVVIGGNLHTCSDKILSAFQCFGRDRDYFVAKNQGTGYGGTLYYQHNETYISLPCKL 725
Db 638 LGYLFGISHTSLDDHSCFLAAGQLLKSSDSFITSSTTSYATVQAQLATSLMKI---- 693
Qy 726 RPCSLSYVTEIPVLPFSGNLSYTHTDNLTKYTYPTVK-GSWGNDSPALFPGGRAPIC 784
Db 694 -----SAQACYNESIHLEKTKYRFSKGFSGWSHVSVAVSVEVCASIPIV 737
Qy 785 LDESALFEQYMPFMKLFQVYAHQEGKEQTEAREFGSSRLVNLALPIGIRPDKESDCOD 844
Db 738 SNGSGLFSSFSISFKLQGFSGTQDGFEESSGEIRSFSSASFRNISLPIGITFEKKSQ-KT 796
Qy 845 ATYNLTIG-YTVDLVRNPDCTTTLRISGDSWKTFTGNLARQALVLRAGNHFCFNSFEA 903
Db 797 RTYYFLGAYIQDLKRDVESGPVLLKNVAVSWDAPMANLDSRAYMFLTNQRALH-RLQT 855
Qy 904 FSQFSFELGRSSRNYNVDLCAKYOF 928
Db 856 LLNVSVCLRQSHSYSLDLGTTYRF 880

RESULT 14
US-09-841-132-189
; Sequence 189, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 189
; LENGTH: 866
; TYPE: PRT
; ORGANISM: Chlamydia
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(866)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-841-132-189

Query Match 17.9%; Score 856; DB 9; Length 866;
Best Local Similarity 30.1%; Pred. No. 2.1e-55;
Matches 260; Conservative 123; Mismatches 369; Indels 114; Gaps 24;
Qy 118 TFIGFSSLSFIASPGSSITTKGAVS-----CST-----GSLSLTKN 154
Db 62 SYCFVSKLHTDPKALFKKGDLSIQNFRFLFTDCSSKSSPSIITHQNGQLSRNN 121
Qy 155 VSLFSLKFNFTDNGGAIKTAKLSLTGTTMSALFSENTSKKGAIQTSALITGNGQEV 214
Db 122 GMSFCRNHAEKGGGAIADAFSLQHNLYLFTAFENSKNGGAIQ-AQTFSLSRNVSPI 180
Qy 215 SPSDNTSSDGAIAITEASVTISNNAKVSFDINKVTGASSSTTGMSCGAICA---YKTS 271
Db 181 SPARNRADLNGAICCSNLICSGNVPNPLFTGNSAT-----NGXICCSIDLNTS 230
Qy 272 TDTKVTLCNOMLLFSNNTSTTAGGAIYVKKLEASGLTLFNRSNVNGTAPKGAIAI 331
Db 231 EKGSLSLACNQTLFASNAKEKGAIAVAKHVLVNGVPVFINNS-----AKIGGAIAI 285
Qy 332 EDGSLSLADSAGDIVFLGNTVTTSTPGTNRSSIDLGTSAKNTALRSAGRIAYPDIT 391
Db 286 QSGGSLILAGSGVLFNQNSQRTSDQGLVRNAIYLEKDAILSLSEARNG-DILFPDPIV 344

Qy 392 TGSSTTVTDVLKVNTPADSLAQYT-----GNIITFGKLSSETAA 432
Db 345 QESS-----KESPLSSLOASVTPTATASPLVIQTSANRSVIFSSRLSEBEKT 396
Qy 433 DSKNLTSLKLPVTLGGTSLKKGVTIQTQFTQADSRLEMDVGTTLLEPADTSTNNL 492
Db 397 PD-NLTSQLOQPFIELKSGRLVLDRAVLSPSLSDPQALLTMEAGTSLKTSXDLKLTXX 455
Qy 493 VINISIDGAKAKIETKATSKNLTLSGTITLLDPTGFYENHSLRNQSYDILELKASG 552
Db 456 SIPLSHSLDTEKSVTH-----APNLSIQKIFLNSGSDENFYENVLLSKENNI-PLLTLP 510
Qy 553 TWTSTAVTPDPIMGEKFGHYGQGTWGPVW---GTGASTTATFNMTKTGYIENPERIGSL 609
Db 511 KQCSHLHLPDGNLSS--HFGYQGDW-TFSWKDSDEGHSLIA--NWTPKNYVHPERQSTL 565
Qy 610 VPNSLWNAFIDTSSILHYLMETANEGLQGDRAF--WCAGLSNEF--HKDSTKTRRFRHLS 665
Db 566 VANTLWNTYSDMQAVQSMINTTAHG--GAYLFTGWSAVSNLFYVHDSGSKPIDNWHRS 623
Qy 666 GGVVIGGNLHTCSDKILSAFQCFGRDRDYFVAKNQGTGYGGTLYYQHNETYISLPCKL 725
Db 624 LGYLFGISHTSLDDHSCFLAAGQLLKSSDSFITSSTTSYATVQAQLATSLMKI---- 679
Qy 726 RPCSLSYVTEIPVLPFSGNLSYTHTDNLTKYTYPTVK-GSWGNDSPALFPGGRAPIC 784
Db 680 -----SAQACYNESIHLEKTKYRFSKGFSGWSHVSVAVSVEVCASIPIV 723
Qy 785 LDESALFEQYMPFMKLFQVYAHQEGKEQTEAREFGSSRLVNLALPIGIRPDKESDCOD 844
Db 724 SNGSGLFSSFSISFKLQGFSGTQDGFEESSGEIRSFSSASFRNISLPIGITFEKKSQ-KT 782
Qy 845 ATYNLTIG-YTVDLVRNPDCTTTLRISGDSWKTFTGNLARQALVLRAGNHFCFNSFEA 903
Db 783 RTYYFLGAYIQDLKRDVESGPVLLKNVAVSWDAPMANLDSRAYMFLTNQRALH-RLQT 841
Qy 904 FSQFSFELGRSSRNYNVDLCAKYOF 928
Db 842 LLNVSVCLRQSHSYSLDLGTTYRF 866

RESULT 15
US-09-738-269-53
; Sequence 53, Application US/09738269
; Publication No. US20030185848A1
; GENERAL INFORMATION:
; APPLICANT: JOHNSTON, STEPHEN A.
; APPLICANT: STEMKE-HALE, KATHERINE
; APPLICANT: SYKES, KATHRYN F.
; APPLICANT: KALTENBOECK, BERNHARD
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR VACCINATION COMPRISING
; TITLE OF INVENTION: NUCLEIC ACID AND/OR POLYPEPTIDE SEQUENCES OF
; TITLE OF INVENTION: CHLAMYDIA PSITTACI
; FILE REFERENCE: UTSD:659
; CURRENT APPLICATION NUMBER: US/09/738,269
; CURRENT FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 679
; TYPE: PRT
; ORGANISM: Chlamydia psittaci
US-09-738-269-53

Query Match 17.8%; Score 851.5; DB 12; Length 679;
Best Local Similarity 30.9%; Pred. No. 3.2e-55;
Matches 243; Conservative 116; Mismatches 283; Indels 145; Gaps 23;
Qy 1 MKSS-FPKVFESTFAIPLS-MIATEVTLDSSASFDGKNGKNGFVSRSQEDAGTTLFKG 58
Db 6 MKNISYGVLLFSSLATATKLLADASVNIATFNGSTSTFNNKQTDNADGTYILGS 65
Qy 59 NVTLENIPGTGTATTKSCFNNTKGLDFTFTGNNGNLLFQTVDAITVAGAVNSVVDKSTT 118

Db	66	AITFEHLNQLKPA NT -SCFANTAGDLTFTGNRRLLYFNNI -SSTAGGAIAISTTADCKTLT	123
Qy	119	FIGFSSLSFIASPGSSIITTKGAVSCSTGSLSTKGNVILLSKFNSTDNGGAIKTAKTILSL	178
Db	124	ISGALQLIFYMSP -RLATNG-	143
Qy	179	TGTTMSALFSENTSKKGGAIQTSDALJITGNQGEVFSFDMTSSDGGAAIFTEASVTISN	238
Db	144	-----	155
Qy	239	NAKVSFIDNKVGTGASSSTTGDMSGGAICAYKTISTDTKVT-----LTGNQMLLFSNNTST	293
Db	156	NSQSSGLNKSAG-----KGVPICEK-STUDVGATSPTLIRNNGEFTLVGNATIS	205
Qy	294	AGGAIYVKKLBSLGGTLFGRNSVNGCTAPKGGAAIAIEDGSELSDSGSDIVFLGNTV	353
Db	206	SGGAIYAEKMTLSGGVYTKFQSNV-----SYDQGAIAIAPNGEISLSADKGNIVPERNLK	261
Qy	354	TSITPGINRSSIDIGTSAKMTALRSAAAGRAIYFYDPITTGSTTTVDLVKNETPADSAL	413
Db	262	IANKQNT-PNAIHIGDNKAKFLQLRAANNKAIFYDPITTTGS--VADRLIINNSQGE-AS	317
Qy	414	QYTGNIIFTGKLSETEAAADSKNLTSLKLLQPVTISGCTLSLKHGVTLOTQAFTOQADSR	473
Db	318	TYDGAIVFSSLNF-THSPECK--LSSFSQGLTLAAGSLVLEEGVCVQAPSPDRAHQSL	374
Qy	474	EMDVGTTLBPADTSTINNLVINISIDGAKKAKIETKATSKNLTLSGTLITLDPDTGTFYVE	533
Db	375	FMPNPGTKLQATQNISVKNLHLNLRI-AEEPAIYITTTDDASSVDICGPPVWHIDEIFYN	433
Qy	534	NHSLRNQSYDILKAS--GTVITSTATPPPIMGKEKHYGYOGTGWPIVGT-----G	585
Db	434	QTVUANELSVLCLNSSLPHLDNITIDDPVAPIMTLESHRGYGGFW-EISWKEQKPLTFG	492
Qy	586	ASTTAT-----FNTKTYGIP-----NPBRIGSLVPNSLWNAPIDISLHLYMETANEGLQ	636
Db	493	KATIAPNKQMLHIWKPSGYVFPFGSGTGFTTSLVPNSLWNLFLDRFSQQALE--KHAVS	550
Qy	637	GDRAFWCAGLSNPHKOSTKTRRPRHISGGVVTGGNLHTCSDKILSAAPCOLFORDRDY	696
Db	551	SGNGIWISSMTNPSFLQSGTNNNHGFRHKSSGYTAGGKIQTLLQDDIFSVSFSOLFGRSKDP	610
Qy	697	FVAKNOGVYGGTLYYQHNE-----TVISLPCKLRPCSLSVPEIPIVLRSGNLSYTHT	750
Db	611	GSATSKDTFLSGSIYAQHSRRLPLIMRFLAGTSTYRPRLLISIPKNLFINFDVLVSYSDV	670
Qy	751	DNDLJTK	757
Db	671	SNHMKVO	677

Search completed: December 16, 2003, 10:59:53
Job time : 40 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2003, 10:49:50 ; Search time 27 Seconds
(without alignments)
3305.353 Million cell updates/sec

Title: US-09-428-122-2
Perfect score: 4774
Sequence: 1 MKSSPPKVFVSTFAIFPLSM.....FELGSSRNYYNDLGAKYQF 928

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3991	83.6	772	H86492	Pmp_3 [imported]
2	2058	43.1	928	G8546	polymorphic outer
3	2058	43.1	928	G81591	polymorphic membra
4	1982	41.5	928	B86546	polymorphic outer
5	1982	41.5	928	B72077	polymorphic membra
6	1965	41.2	936	C72078	polymorphic outer
7	1964	41.1	936	C8546	polymorphic outer
8	1964	41.1	936	B81591	polymorphic membra
9	1940	40.6	930	D8546	polymorphic outer
10	1940	40.6	930	A81591	polymorphic membra
11	1936	40.6	930	D72078	polymorphic outer
12	1855	38.9	928	H86546	polymorphic outer
13	1855	38.9	928	D72077	polymorphic outer
14	1855	38.9	949	F81591	polymorphic membra
15	1573	32.9	841	E72130	polymorphic outer
16	1444	30.2	1276	B86546	polymorphic membra
17	1444	30.2	1276	C81591	polymorphic outer
18	1442.5	30.2	922	B72131	polymorphic membra
19	1442.5	30.2	922	E86491	polymorphic outer
20	1441.5	30.2	922	F81539	polymorphic membra
21	1383.5	29.0	1407	B72078	polymorphic outer
22	1377.5	28.9	973	B86547	polymorphic membra
23	1377.5	28.9	973	F72076	polymorphic outer
24	1377.5	28.9	995	C81593	polymorphic membra
25	1246	26.1	712	B86492	polymorphic outer
26	1120.5	23.5	1013	G71460	probable outer mem
27	1051	22.0	987	H81722	polymorphic membra
28	1017.5	21.3	445	B86493	Pmp_5 [imported]
29	916	19.2	359	C86493	Pmp_4 [imported]

30	911	19.1	867	2	F81721	polymorphic membra
31	863	18.1	878	2	B71460	probable outer mem
32	821	17.2	494	2	D86493	polymorphic outer
33	792	16.6	427	2	A86493	polymorphic outer
34	768	16.1	186	2	G86492	polymorphic outer
35	685.5	14.4	1609	2	A86611	probable outer mem
36	685.5	14.4	1609	2	H72013	polymorphic membra
37	672.5	14.1	978	2	G72076	polymorphic outer
38	670	14.0	947	2	G86557	polymorphic membra
39	667.5	14.0	978	2	B81593	polymorphic membra
40	667.5	14.0	978	2	C86547	polymorphic outer
41	666	14.0	947	2	D72067	polymorphic membra
42	666	14.0	1723	2	H86557	polymorphic membra
43	666	14.0	1723	2	E72067	polymorphic membra
44	666	14.0	1732	2	C81601	polymorphic membra
45	665.5	13.9	946	2	C86549	polymorphic outer

RESULT 1
H86492
Pmp_3 [imported] - Chlamydothila pneumoniae (strain J138)
C:Species: Chlamydothila pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C:Accession: H86492
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: H86492
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-772 <STO>
A:Cross-references: GB:BA000008; NID:g8978389; PIDN:BAA98226.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: pmp_3_2

Query Match 83.6%; Score 3991; DB 2; Length 772;
Best Local Similarity 99.9%; Pred. No. 2.3e-208;
Matches 771; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	157	LLFSKNFSTNDGGAITAKTSLTGTMTSALFSENSSKKGAIQTSALTITGNQGEVSF	216
Db	1	MLFSKNFSTNDGGAITAKTSLTGTMTSALFSENSSKKGAIQTSALTITGNQGEVSF	60
Qy	217	SNTSSDSGAAIFTEASVTISNNAKVSFDNKTGASSTTGDMSGGAIQTSALTITGNQGEVSF	276
Db	61	SNTSSDSGAAIFTEASVTISNNAKVSFDNKTGASSTTGDMSGGAIQTSALTITGNQGEVSF	120
Qy	277	TLTGNQMLFSSNNTTGTAGGAIYVKKLELASGGLTLFSENSSKKGAIQTSALTITGNQGEVSF	336
Db	121	TLTGNQMLFSSNNTTGTAGGAIYVKKLELASGGLTLFSENSSKKGAIQTSALTITGNQGEVSF	180
Qy	337	LSLSADSGDIVLGNVTSTTGTNRSSIDLTSAMTALRSAAAGRAIYFYDPIITGSST	396
Db	181	LSLSADSGDIVLGNVTSTTGTNRSSIDLTSAMTALRSAAAGRAIYFYDPIITGSST	240
Qy	397	TVTDVLKVNETPADSALQVTGNIIFTGEKLSATEAASDKNLSKLLQPVTLSSGGLSLKH	456
Db	241	TVTDVLKVNETPADSALQVTGNIIFTGEKLSATEAASDKNLSKLLQPVTLSSGGLSLKH	300
Qy	457	GVTLTQTAFTQADSRLMDVGTTLTLEPADTSTINNLVINISSIDGAKKAKIETKATSKNL	516
Db	301	GVTLTQTAFTQADSRLMDVGTTLTLEPADTSTINNLVINISSIDGAKKAKIETKATSKNL	360
Qy	517	TLSGTITLLDPTGTTFYENHSLRNPOSIDILEKASGTVTSTAVTTPDIMGKHFHYGQQT	576
Db	361	TLSGTITLLDPTGTTFYENHSLRNPOSIDILEKASGTVTSTAVTTPDIMGKHFHYGQQT	420
Qy	577	WGPIVMGTGASTATFNNKTGYIPNPERIGSLVPSNLWNAFIDISSLHYLMETANEGLQ	636

ALIGNMENTS


```
Db 421 WGVWGTGASTATATFNWTKGYPNPERIGSLVPSNLWNAFIDISLHVLMTANEGLQ 480
QY 637 GDRAFWCAGLSNPFHKDSTKTRGRFRLHSGGYVIGNHLTCSDKILSAAPCOLFGRDRY 696
Db 481 GDRAFWCAGLSNPFHKDSTKTRGRFRLHSGGYVIGNHLTCSDKILSAAPCOLFGRDRY 540
QY 697 FVAKNGQTVVGGTLYYOHNETYISLPCKLKRPCSLSVYPTPIPLVFSGNLSYTHTDNDLKT 756
Db 541 FVAKNGQTVVGGTLYYOHNETYISLPCKLKRPCSLSVYPTPIPLVFSGNLSYTHTDNDLKT 600
QY 757 KYTTPYTVKSGWNGDSFALFEGGRAPICLDESALFEQYMPFMKLQFYVAHQEGFKEQGT 816
Db 601 KYTTPYTVKSGWNGDSFALFEGGRAPICLDESALFEQYMPFMKLQFYVAHQEGFKEQGT 660
QY 817 AREFGSRLVNLAIPGIRDFKESDCQDATYNTLGLYVDLVRNPDCTTTLRISGDSWK 876
Db 661 AREFGSRLVNLAIPGIRDFKESDCQDATYNTLGLYVDLVRNPDCTTTLRISGDSWK 720
QY 877 TFGTNLARQALVLRAGNHFCFNSNFEAFSOFSELGSSRNYNVDLGAKYQF 928
Db 721 TFGTNLARQALVLRAGNHFCFNSNFEAFSOFSELGSSRNYNVDLGAKYQF 772

RESULT 2
G86546
polymorphic outer membrane protein G family [imported] - Chlamydomophila pneumoniae (strain
C;Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 15-Jun-2001
C;Accession: G86546
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ie
Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349; PMID:10871362
A;Accession: G86546
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-928 <STO>
A;Cross-references: GB:BA000008; NID:g8978821; PIDN:BA98657.1; GSPDB:GN00142
A;Experimental source: strain J138
C;Genetics:
A;Gene: pmp 10
C;Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Query Match 43.1%; Score 2058; DB 2; Length 928;
Best Local Similarity 46.7%; Pred. No. 1.2e-103;
Matches 444; Conservative 142; Mismatches 320; Indels 44; Gaps 20;

QY 1 MKSFPKFPV-STFAIF---PLSMIATETVLDSSASFDGNKN-GNFSVRSEQEDAGTYL 55
Db 1 MKSQFSLVLSSTLACFTSCSTVFAAENIGPDSFDGNTGTTPKNT--TTGIDYT 58

QY 56 FKGNTVLENIPTGTATKSCFNNTKGDLTFTGNGNSLLFQTVDAGTVAAGAVNSVVDK 115
Db 59 LTGDDITLQNL-GDSAAITKGCFSDDTESLFAKGYSLFLNKS-SAEGAAL-SVTTDK 115

QY 116 STTFIFGSSLSFIASPGSSITTT--GKGVSCSTGSLTKNVLSLFSKFNSTNGGAITA 173
Db 116 NLSLTGSSSTFLAAPSIVITPSGKGVKCG-GDUTFDNNNGTILFKQDYCEBNGAIST 174

QY 174 KTLISLTGTTMSALFSENSTSS---KKGAIQTSDALITTNQGEVFSFSDNTSSDSGAAIFT 230
Db 175 KNLISLKNSTGSIPEGNKSSATGKGGAI CATGTVDITNTNATPLFSNNTAEEAGGAINS 234

QY 231 EASVTISNNAKVPIDNKVTTGASSTTGDMSGGAICAYKSTDTKVTLTGQMLLSNNNT 290
Db 235 TGNCTITGNTSLVSENSVT---ATAG--NGGAL-----SGDADVTISGNQSVTFSGNQ 283

QY 291 STTAGGAIYVKKLELAS--GGLTLFNSVNGGTAPKGGAIADSDGSLSDSGDIVF 348
Db 284 AVANGGAIYAKKTLASGGGGGSIFFSNIVQGTAGNGAISILAAEGCSLSAEAGDITF 343

QY 349 LGNTVITSTTP-GTNRSSIDLGTSAKMTALRSAGRAIYFYDPIITGSSTTVTDVLKNET 407
```

```
Db 344 NGNAIVATTPTQTKRNSIDIGSTAKITNLRALSGBHSIFFIDPITANTAADSTDTLNLKA 403
QY 408 PADSALOYTGNITIFTGEKLSSEAAADSKNLTSLKLPQVTLSGGTLSLKHGVTLTQQAFTQ 467
Db 404 DAGNSTDYSGSIVFSGEKLSEDEAKVADNLSTLQKQPVTLTLAGNLVLKGVTLDTKGFTQ 463
QY 468 QADSRLEMDVGTLEPA-DTSTINNVLVINISIDGAKKAKIETKATSKNLTJSGTITLLD 526
Db 464 TAGSSVINDAGTTLKASTEVEVTLTGLSIFVDSLGEKGVVIAAASAKNVALSGPILLD 523
QY 527 PTGTFVNHSLRNPOSYDITLELKASQTVTSTAVTPDPIMGKFRPHYGYQCTWGPVW---- 582
Db 524 NQGNAYENHDLKTQDFSFVQLSALGTATTDVPAVPTVATPTHGYQGTWG-MTWVDDT 582
QY 583 -GTGASTTATFNWTKGYIPNPERIGSLVPSNLWNAFIDISLHVLMTANEGLQDRAF 641
Db 583 ASTPKTKTATLAWNTGYLNPBERQGPLVPSNLGWSFSDIQAIOGVIERALSITLCSDRGF 642
QY 642 WCAGLSNPFHKDSTKTRGRFRLHSGGYVIGNHLTCSDKILSAAPCOLFGRDRYFVAKN 701
Db 643 WAAGVANFLDKDKGKRYRHSKGGYATGGAATCSENLIISFAFCQLFGSDKDFLVAKN 702
QY 702 QGTVYGGTLYYOHNETYISLPCKLKRPCSLSVYF---TEIPVLFSGNLSYTHTDNDLTKY 758
Db 703 HDTYAGAFYIQH---ITECSGFICLLDKLPGWSHKLPLVLEQLAYSHVNSNDLTKY 758
QY 759 TYTTPYTVKSGWNGDSFALFEGGRAPICLDESALFEQYMPFMKLQFYVAHQEGFKEQGT 818
Db 759 TAYPEVKGSGWNGNAFNMGLGASSHSYPEYLHCFDTPYVYIKLNLTVIRQDSFSEKTEGR 818
QY 819 EFGSRLVNLAIPGIRDFKESDCQDATYNTLGLYVDLVRNPDCTTTLRISGDSWK 878
Db 819 SFDDSNLFLNLSIPGVKFEKFSDCNDFSYDLTSLVYVPDLIRNDPKCTTALVISGASWET 878
QY 879 GTNLARQALVLRAGNHFCFNSNFEAFSOFSELGSSRNYNVDLGAKYQF 928
Db 879 ANNLRARQALQVRAGSHYAFSPMFEVLQGFVFGSSRNYNVDLGKQFQF 928

RESULT 3
G81591
polymorphic membrane protein G family CP0303 [imported] - Chlamydomophila pneumoniae (
C;Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 15-Jun-2001
C;Accession: G81591
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hick
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR3
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: G81591
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-928 <REA>
A;Cross-references: GB:AE002192; GB:AE002161; NID:g7189226; PIDN:AAF38160.1; PID:g71
A;Experimental source: strain AR39, HL cells
C;Genetics:
C;Gene: CP0303
C;Superfamily: Chlamydomophila pneumoniae polymorphic outer membrane protein G

Query Match 43.1%; Score 2058; DB 2; Length 928;
Best Local Similarity 46.7%; Pred. No. 1.2e-103;
Matches 444; Conservative 142; Mismatches 320; Indels 44; Gaps 20;

QY 1 MKSFPKFPV-STFAIF---PLSMIATETVLDSSASFDGNKN-GNFSVRSEQEDAGTYL 55
Db 1 MKSQFSLVLSSTLACFTSCSTVFAAENIGPDSFDGNTGTTPKNT--TTGIDYT 58

QY 56 FKGNTVLENIPTGTATKSCFNNTKGDLTFTGNGNSLLFQTVDAGTVAAGAVNSVVDK 115
Db 59 LTGDDITLQNL-GDSAAITKGCFSDDTESLFAKGYSLFLNKS-SAEGAAL-SVTTDK 115

QY 116 STTFIFGSSLSFIASPGSSITTT--GKGVSCSTGSLTKNVLSLFSKFNSTNGGAITA 173
Db 116 NLSLTGSSSTFLAAPSIVITPSGKGVKCG-GDUTFDNNNGTILFKQDYCEBNGAIST 174

QY 174 KTLISLTGTTMSALFSENSTSS---KKGAIQTSDALITTNQGEVFSFSDNTSSDSGAAIFT 230
Db 175 KNLISLKNSTGSIPEGNKSSATGKGGAI CATGTVDITNTNATPLFSNNTAEEAGGAINS 234

QY 231 EASVTISNNAKVPIDNKVTTGASSTTGDMSGGAICAYKSTDTKVTLTGQMLLSNNNT 290
Db 235 TGNCTITGNTSLVSENSVT---ATAG--NGGAL-----SGDADVTISGNQSVTFSGNQ 283

QY 291 STTAGGAIYVKKLELAS--GGLTLFNSVNGGTAPKGGAIADSDGSLSDSGDIVF 348
Db 284 AVANGGAIYAKKTLASGGGGGSIFFSNIVQGTAGNGAISILAAEGCSLSAEAGDITF 343

QY 349 LGNTVITSTTP-GTNRSSIDLGTSAKMTALRSAGRAIYFYDPIITGSSTTVTDVLKNET 407
```



```
QY 590 --ATFNWTKGYIPNPERIGSLVPSNLWNAFIDISSLHYLMETANEGLQDRAPWCAGLS 647
Db 602 KEATATWTKGTFVPSPERKSALVCNTLWGVFTDIRSLQQLVEIGATGMEHKGFWSSMT 661
QY 648 NFPHKDKSTKRRGRFRLHSGGYVIGNHLHTCSDKILSAAPCOLFGRDRDYFAKNQGVYV 707
Db 662 NFLHKGTGDNKRKFRHTSGGYVIGGSAHTPKDDLFTFAPCHLPARKDCFIANNRTYG 721
QY 708 GTLYYQHNET-----YISL-PCRLPCSLSYVTEIPVLSFGNLSYTHTDNLKTYTY 761
Db 722 GTLFFKSHSTLQPNVYLRGLGRAKFSESAIEKFPREIPLALDVQVSPSHSDNRMETHYTS 781
QY 762 PTYKSGWGNDSFALEFCGRAPICL-DESALFEQYMPMKLFYVAHQEGKEQGTAREF 820
Db 782 PESEGSNSNECIAGGIGLDLPVLSNPHLPKTFIPQMKVEMYVYVQNSFFESSDGRGF 841
QY 821 GSSRLVNLALPIGIRFKESDCODA-TYNLTGLVTVDLVRSNPDCTTTLRISGDSWKTG 879
Db 842 SIGRLNLSIFVGAKEF-VQGDIGDSYTDLSGFFVSDVYRNFPQSTATLWSPDSWKIRG 900
QY 880 TNLARQALVLRAGNHCFNENFAFSQFSELPFGSSRNYNVDLGAKYQF 928
Db 901 GNLSRQAFLLRGSNNYVYNSNCELFCHYAMELRGSSRNYNVDVGTKLRF 949

RESULT 15
E72130
polymorphic membrane protein G family CP0761 [imported] - Chlamydothila pneumoniae (stra
C;Species: Chlamydothila pneumoniae, Chlamydia pneumoniae
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 15-Jun-2001
C;Accession: E72130; G81541
R;Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A;Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606; PMID:10192388
A;Accession: E72130
A;Molecule type: DNA
A;Residues: 1-841 <ARN>
A;Cross-references: GB:AE001363; NID:94376263; PIDN:AA18172.1; PID:9437627
A;Experimental source: strain CWL029
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: G81541
A;Molecule type: DNA
A;Residues: 1-841 <REA>
A;Cross-references: GB:AE002235; GB:AE002161; NID:g7189672; PIDN:AAF38561.1; PID:g718967
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: pmp2; CP0761
C;Superfamily: Chlamydothila pneumoniae polymorphic outer membrane protein G

Query Match 32.9%; Score 1573; DB 2; Length 841;
Best Local Similarity 39.4%; Pred.No.1.8e-77;
Matches 371; Conservative 121; Mismatches 310; Indels 140; Gaps 20;

QY 12 TFAIFPLSMIATETVLDSSAPDG-NKNGNFSVRESQEDAGTTFYFKGNVTLENIPGTGT 70
Db 15 TLSMNLGAATTEELSGNSPDGTTSTTSFSKTSATDGTNYVFKDSVVIENVPKGE 74
QY 71 AITKSCFNN--TKGDLFTFGNLSLLPQTVDAGTGAAGAVNSVVDKSTFTFGFSLSFI 128
Db 75 TQSTSCFKNDAAAGDLNFGGFSFTFSNIDATTASGAIGSEAAKNTVTLSGFSALSFL 134
QY 129 ASFGSSITTKGKAVSCSTGSLSTKNVSLFLSKNFTDNGGAIKTAKTLSITGTMALFS 188
Db 135 KSPASIVTNGLGAINVK-GNLSLDDNDKVLIQNFSTGDDGAINC----- 178
QY 189 ENTSKKGKGIQTSALDITITNGQGVFSFSDNTSSDGAIFTBEASVTISNNAKVSFIDNK 248
```

```
Db 179 -----AGSLKIANNKLSLSFI--- 193
QY 249 VTGASSSTTGDMSGGCAICAYKTSTDTKVTLTGNQMLLFNSNNTSTTAGGAIYVKKLELASG 308
Db 194 -----GNSSSTGGAIHTKNTLSSG 214
QY 309 GLTLFSRNSVNGGTAP-----KGGATAIEDSGELSLSADSGLIVFLGNVTSTTPTGT-NRS 363
Db 215 GETLF-----QGNAPTAAAGKGAIAIADSGTSLISGSDGDIIFEGNTIGAT--GTVSHS 267
QY 364 SIDLGTSAKWTALRSAGRAIFYDPITTTGSSSTTVTVLKVNETPADSALQYTGNTIIFTG 423
Db 268 AIDLGTSAKITALRAAQHTIIFYDITVTGTSVADALNINSPTGDNKEYTGTIIVFSG 327
QY 424 EKLSETEAADSNNLSKLLQPVTLSGGTLSLKHGVTLOTQAFTOQADSLEMDVGTTLRP 483
Db 328 EKLTEAEAKDEKNTSKLLQNVAFKNGTVVLKGDVLSANGFSQDANSKILMDLGTSL-V 386
QY 484 ADTSTI--NNLVINISSIDGAKAKIETKATSKNLTSLGTITILLDPTGTFFYENHSLRNQ 541
Db 387 ANTESIELTNLEINIDSLRNGKKIKLSAATAQKDIRIDRPVLAISDESFYQNGFLNEIDH 446
QY 542 SYD-ILELKASGTVTSTAVTPDPIMGEXPHYGVQGTWGPVAGTGASTTATENWTKGYI 600
Db 447 SYDGIILELDAGKDIVISADRS-IDAVQSPYGYQGM-TINWSTD-DKATVSWAKQSPN 503
QY 601 PNPERIGSLVPSNLWNAFIDISSLHYLMETANEGLQDRAPWCAGLSNPFPHKDKSTKTRRG 660
Db 504 PTAEQEAPLVPNLLMGSTFIDVRSFQNFIELGTGAPYEKRFVWAGISNVLHRSRENQK 563
QY 661 FRHLGGYVIGNHLHTCSDKILSAAPCOLFGRDRDYFAKNQGVYGGTLYYQHNETYIS 720
Db 564 FRHSGGAVVGASTRMPPGDDTSLGFAQLFARDKDYFMNTNFAKTYAGSLRLQHDASLYS 623
QY 721 LFC-----KLRPCSLSYVTEIPVLSFGNLSYTHTDNLKTKY-----TTVPTVKGSWG 769
Db 624 VVSILLGEGGLEILLPVYSKTLPCSFYQLSYGHYTHRMKTESLP PPPPTLSTHTSWG 683
QY 770 NDSFALFEGGRAPICLDESA---LFEQYMPFMKLFQVVAHQEGKEQGTAREFGSSRLV 826
Db 684 GYVWAGELGTR--VAVENTSGRGFFQEVTFPVKVQAVYARQDSFVELGAI SRDSDSHLY 741
QY 827 NLALPIGIFPDKESQDQATNLTGTYVDLVRSPDCTTTLRISGDSWKTFTGTNLARQA 886
Db 742 NLALPIGLIKLEKR--FABQYHVAVMSPDVCNPKCTTTLLSNQGSNKTGSLARQA 799
QY 887 LVLRAGNHCFNENFAFSQFSELPFGSSRNYNVDLGAKYQF 928
Db 800 GIVQASGFFSLGAAAEELFGNFGFWRGSSRSYNVDAGSKIKF 841

Search completed: December 16, 2003, 10:54:43
Job time : 30 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2003, 10:34:04 ; Search time 17 Seconds

(without alignments)

2567.105 Million cell updates/sec

Title: US-09-428-122-2

Perfect score: 4774

Sequence: 1 MKSFPKFFVSTAIPLSM.....FELRGSSRNVDLGAKYQF 928

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2058	43.1	928	1	PM10_CHLPP
2	1982	41.5	928	1	PM10_CHLPP
3	1964	41.1	936	1	PM10_CHLPP
4	1936	40.6	930	1	PM10_CHLPP
5	1855	38.9	928	1	PM10_CHLPP
6	1573	32.9	841	1	PM10_CHLPP
7	1444	30.2	1276	1	PM10_CHLPP
8	1442.5	30.2	922	1	PM10_CHLPP
9	1377.5	28.9	973	1	PM10_CHLPP
10	1120.5	23.5	1013	1	PM10_CHLPP
11	1051	22.0	987	1	PM10_CHLPP
12	911	19.1	867	1	PM10_CHLPP
13	863	18.1	878	1	PM10_CHLPP
14	685.5	14.4	1609	1	PM10_CHLPP
15	667.5	14.0	978	1	PM10_CHLPP
16	666	14.0	947	1	PM10_CHLPP
17	666	14.0	1723	1	PM10_CHLPP
18	665.5	13.9	946	1	PM10_CHLPP
19	647	13.6	952	1	PM10_CHLPP
20	603.5	12.6	938	1	PM10_CHLPP
21	603	12.6	975	1	PM10_CHLPP
22	597	12.5	1016	1	PM10_CHLPP
23	594	12.4	980	1	PM10_CHLPP
24	587.5	12.3	976	1	PM10_CHLPP
25	576.5	12.1	1531	1	PM10_CHLPP
26	559	11.7	1754	1	PM10_CHLPP
27	549	11.5	976	1	PM10_CHLPP
28	548.5	11.5	1672	1	PM10_CHLPP
29	544	11.4	1520	1	PM10_CHLPP
30	541.5	11.3	964	1	PM10_CHLPP
31	533.5	11.2	1460	1	PM10_CHLPP
32	523	11.0	1770	1	PM10_CHLPP
33	490	10.3	1034	1	PM10_CHLPP

34 475.5 10.0 514 1 PM12_CHLPP
35 453 9.5 1025 1 PM12_CHLPP
36 269 5.6 2249 1 OMPA_RICRI
37 251 5.3 2021 1 OMPA_RICRI
38 238.5 5.0 1025 1 SLAP_CAUCR
39 237.5 5.0 881 1 PRY3_YEAST
40 234 4.9 1645 1 OMPB_RICRY
41 232 4.9 1643 1 OMPB_RICPR
42 230 4.8 995 1 YIQ9_YEAST
43 230 4.8 1250 1 YFAL_ECOLI
44 226.5 4.7 1609 1 FIG2_YEAST
45 224.5 4.7 2003 1 YDBA_ECOLI

ALIGNMENTS

RESULT 1

PM10_CHLPP STANDARD; Q9RB64; PRT; 928 AA.
AC Q9RB65; Q86163; Q9RB64; Q9S6P2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable outer membrane protein pmp10 precursor (Polymorphic membrane protein 10) (Outer membrane protein 5).
DE PM10 OR OMP5 OR CP0303.
GN Chlamydia pneumoniae (Chlamydia pneumoniae).
OS Bacteria; Chlamydiales; Chlamydiaceae; Chlamydiales.
OC NCBI_TaxID=83558;
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VR1310;
RX MEDLINE=20007584; PubMed=10539856;
RA Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT "Molecular biology of Chlamydia pneumoniae surface proteins and their role in immunopathogenicity.";
RL Am. Heart J. 138:S491-S495(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., Bass S., White O., Hickey E.K., Peterson J., Utterback T., Berry K., Dodson R., Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R., Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K., Shiba T., Ishii K., Hattori M., Kohara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CML029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [4]
RP SEQUENCE OF 1-914 FROM N.A.
RC STRAIN=CML029/VR-1310;
RX MEDLINE=99081766; PubMed=9864239;
RA Knudsen K., Madsen A.S., Mygind P., Christiansen G., Birkelund S.;
RT "Identification of two novel genes encoding 97- to 99-kilodalton outer membrane proteins of Chlamydia pneumoniae.";
RL Infect. Immun. 67:375-383(1999).
CC -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES) (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; AJ13034; CAB37071.1; -
 EMBL; AE002192; AAF38160.1; -
 EMBL; AP002546; BAA98657.1; -
 EMBL; AJ001311; CAA04671.1; -
 PIR; G81591; G81591.
 PIR; G86546; G86546.
 PHCI-2DPAGE; O86163; -
 TIGR; CP0303; -
 InterPro; IPR006315; Autotransport.
 InterPro; IPR003368; Chlamydia_PMP.
 Pfam; PF02415; DUF145; 2.
 TIGRFAMS; TIGR01414; autotrans_bar1; 1.
 TIGRFAMS; TIGR01376; POMP repeat; 6.
 Outer membrane; Signal; Multigene family.
 SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 928 PROBABLE OUTER MEMBRANE PROTEIN PMP10.
 SQ SEQUENCE 928 AA; 97229 MW; 0590D5206A1DD0E1 CRC64;

Query Match 43.1%; Score 2058; DB 1; Length 928;
 Best Local Similarity 46.7%; Pred. No. 1.2e-104;
 Matches 444; Conservative 144; Mismatches 320; Indels 44; Gaps 20;

QY 1 MKSFPKVF-STFAIF---PLSMIATETVLDSASPDGNKN-GNFSVRESQDAGTYL 55
 Db 1 MKSQFWLUSLTLACTSCSTVPAFAENIGSDSFDGSTNTGTTPKNT--TTGDIYT 58
 QY 56 FKNVTLENIPGTGTAITKSCFNNTKGLFTFTGNKSLLPQTVDAGTVAGAAVNSSVDK 115
 Db 59 LTGDITLQNL-GPSAALTGCFSDTTLSLFAKGYLSLFLNKS-SAEGAAL-SVTYDK 115
 QY 116 STTFIFSSLSFTASPGSSITT--GKAVSCSTGSLSLTKVNLFSKNFSTDNGAITA 173
 Db 116 NLSLTGFSLSLTFLAAPSSTVITPSGKAVKCG-GDLTFDNNITLFPKDYCEENGGAIST 174
 QY 174 KTLSLAGTMSALFSENTSS--KKGAIQTSALAITGNGEVSFSDNTSSDGAIFT 230
 Db 175 KNLKSNSTGSISEFGNKSATGKGAICATGTVDTNNATPLFSNNIAEAGGAINS 234
 QY 231 EASVTISNNAKVFIDNKVTGASSSTTGMDSGGAICAYKTSITDKVTLTGNOMLFSNNT 290
 Db 235 TGNCTITGNTSLFSENSVI---ATAG--NGCAL-----SGDADVTISGNQSVTSGNQ 283
 QY 291 STTAGGAIYKKLELAS--GGLTLPFSNVNGTAPKGAIAIEDSGELSLSDSDIVP 348
 Db 284 AVANGGAIYAKKLTLAGGGGGISFNNIVQGTAGNGGAISILAAEGCSLSAEGDITF 343
 QY 349 LGNTVSTTTP-GNRRSIDLGTSAKMTALRSAGRAIFYDPIITGSSITVTDLVKNET 407
 Db 344 NGNAIVATTPQTKRNSIDIGSTAKITNLRAISGHSIFFYDPIATANTADSTPLNKA 403
 QY 408 PADSAIQTGNIIFTGKLESEAAADSKNLTSLKLPVTLSSGGLTSLKHGVITQTQAF 467
 Db 404 DAGNSTDYSGSIVFSGEKLSDEAKVADNLITSLKQPVTLTGNLVLKRGVILDTKGFTQ 463
 QY 468 QADSRLEMDVGTLEPA-DTSTINNLVINISSIDGAKKAKIEPKATSKNLTLSGTTLLD 526
 Db 464 TAGSSVIMDAGTTLKASTEEVITGLSIPVDSLGEKGVVIAAASAKNVALSGPILLD 523
 QY 527 PTGTFENHSLRNPQSDVILELKASGTVSTATVTPDPMGEKHYGQGTGPIVW---- 582
 Db 524 NQGNALENHDLGKTQDFSPQLSALGTATTIDVPAVPTVATPHYGQGTGWTG-MTWVDDT 582
 QY 583 -GTGASTTATFNWTKTGTYIPNPRIGSLVPNSLWNAFIDTSSLHLYMETANEGLQDRAF 641
 Db 583 ASPTPKTATLWNTNGYLPNPRQGPLVNSLWGSFSDIOAIQGVIERALSILCSDRGF 642

QY 642 WCAGLSNFFHKDSTKTRGRHLSGGYVIGNLTGCDKILSAAFQCLGRDRDYFAKN 701
 Db 643 WAAGVANFLDKDKGKRYRKHSGGYAIGGAQTCSNLSFAFQCLGSGDKDFVAKN 702
 QY 702 QGTVYGTLYYQHNETYISLPCKLRPCSLSYVP---TEIPVLFSGNLSYTHDNDLTKY 758
 Db 703 HTDTYAGAFYIQH---ITECSGFIGCLLDKLPFGSMHKEPLVLEGQLAYSHVSNLTKY 758
 QY 759 TTYPTVKGSGWGNDSFALFEGGRAPICLDESALFEQYMPKLOFYVAHQEGKEQGTAR 818
 Db 759 TAYPEVKGSGWGNNAFNMMLGASHSPYVYLHCFDITYPIKLNLTVIRODSFSEKGTGR 818
 QY 819 EFGSSSLVNLALPIGIRPKESDCQDATYNLTGLYTVDLVRSPNCDTTLTIRISGGSWKT 878
 Db 819 SFDDSNLNLSLPIGVKPEKFSDCNDFSYDLTSLVYDPIRNDPKRTTALVLSGASWEY 878
 QY 879 GTNLRQALVLRAGNHCFCNSNFEAFSPQSFELGSSRNYNVDLGAQYQF 928
 Db 879 ANNLARQALQVRAGSHVAFSPMFEVLGQFVFEVRGSSRYINVDLGGKQF 928
 RESULT 2
 PMP9_CHLPN
 ID_PMP9_CHLPN STANDARD; PRT; 928 AA.
 AC Q92398;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable outer membrane protein pmp9 precursor (Polymorphic membrane
 DE protein 9) (Outer membrane protein 10).
 GN PMP9 OR OMP10 OR CPN0447 OR CP0306.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 ON NCBI_TaxID=83558;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VR1310;
 RX MEDLINE=20007584; PubMed=10539856;
 RA Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
 RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;
 RT "Molecular biology of Chlamydia pneumoniae surface proteins and their
 RT role in immunopathogenicity.";
 RL Am. Heart J. 138:S491-S495(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CWL029;
 RX MEDLINE=99206506; PubMed=10192388;
 RA Kalman S., Mitchell W., Marathe R., Lammel C., Pan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
 RL Nat. Genet. 21:385-389(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AR39;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey B.K., Peterson J., Utterback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
 RT pneumoniae AR39.";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J138;
 RX MEDLINE=20330349; PubMed=10871362;
 RA Shitai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 RT from Japan and CWL029 from USA.";
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)

RP SEQUENCE FROM N.A.
 RC STRAIN=CWL029;
 RX MEDLINE=99206606; PubMed=10192388;
 RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 RT "Comparative Genomes of Chlamydia pneumoniae and C. trachomatis.";
 RL Nat. Genet. 21:385-389(1999).
 RN [4].
 RP SEQUENCE FROM N.A.
 RC STRAIN=AR39;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
 RT pneumoniae AR39.";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 RN [5].
 RP SEQUENCE FROM N.A.
 RC STRAIN=J138;
 RX MEDLINE=20330349; PubMed=10871362;
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 RT from Japan and CWL029 from USA.";
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -! SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
 CC (POTENTIAL).
 CC -! SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AJ001311; CRA04672.1; -;
 DR EMBL; AJ133034; CAB37072.1; -;
 DR EMBL; AF001628; AAD18593.1; -;
 DR EMBL; AF002192; AAF38159.1; ALT_INIT.
 DR EMBL; AP002546; BAA98658.1; -;
 DR EMBL; D72077; D72077.
 DR F01; H86546; H86546.
 DR PHCI-2DPAGE; O86164; -;
 DR TIGR; CP0302; -;
 DR InterPro; IPR003368; Chlamydia_PMP.
 DR Pfam; PF02415; DUF145.2.
 DR TIGRPFAM; TIGR01376; POMP_repeat; 6.
 KW Outer membrane; Signal; Multigene family; Complete proteome.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 928 PROBABLE OUTER MEMBRANE PROTEIN PMP11.
 SQ SEQUENCE 928 AA; 98903 MW; 788BCDD62C911402 CRC64;

Query Match 38.9%; Score 1855; DB 1; Length 928;
 Best Local Similarity 42.7%; Pred. No. 1.3e-93;
 Matches 405; Conservative 171; Mismatches 331; Indels 42; Gaps 20;
 QY 1 MKSFFKFFVFSTPAIFP--LSMIATETVLDSSASPDGN-KNGNFSVRESQEDAGTTLK 57
 DB 1 MKTISFWLVSSVLAISCHLOSLANEELSPDSDSFGNIDSGTFTPTKTS-----ATTYSUT 56
 QY 58 GNVLTENIFGTGTAITKSCFNNTKGLTFTNGNSLLFTQVDAGTVAGAAVNSVVDKST 117
 DB 57 GDVFFVE-FKGTPLSDSCEKQTDDNLTLFLNGHSLTFGFDAGTHAGAAA-STTANKML 114
 QY 118 TFIGFSSLSFIAPSGSITTKGAVSCSTGSLTKRNVSLFSPKFNSTNGGGAITAKTILS 177
 DB 115 TFGSGLLSFDSPSPSTVTATGQTLIS-SAGGVNLENIRKLVVAGNFSTADGGAIGASFL 173

QY 178 LTGTTMSALPSENTSSKKGAGIQTSDALTITNGQGEVFSFSDNTSSDGAIFTASVTIS 237
 DB 174 LTGTSGDALFSNNSSSTKGAIATAGARIANNNTGVVRFSLNIASTSGGAIDDEGTSILS 233
 QY 238 NNAKVSFIDNKVYTGSSSTTGDMSGCAICAYKTSTDTKVLITGNQMLLSNNNTSTTAGGA 297
 DB 234 NKKFLYP-----EGNAKTT-----GGAICNTKASGSPRLIINNKTLLIFASNAETSGGA 284
 QY 298 IYVKLELASGGGLTLFNRNSVNGGTAPKGAIAIEDSGELSLSDSGSDGVFLGNTVTST- 356
 DB 285 IHAKKLALSSGGFTFLRNNVSSAT-PKGALSIDASGELSLSAETGNITFVNTLITTTG 343
 QY 357 -TPGWRSSIDIGTSAKMTALSAAGRAIYFVDPIPTGTSSTTVTLVKVNETPADSALOY 415
 DB 344 STDTPKRNAINIGSNGKFTELRAAKNHTIFFYDPIT--SEGTSSDVLKINNGSAGALNPFY 401
 QY 416 TGNIIPTGKLESEAAADSKNLTSLKLPVTLSSGGTSLSKHGVTLCOTQAFQOASRLWM 475
 DB 402 QGTILFSGETLTADKLKVNLSKSTQPVLSGGKLLQKGVTLLESTFSQEAAGLLGM 461
 QY 476 DVGTTLE-PADTSTINNLVINISSIDGAKKAKIETKATSKNLTLSCTITLLDPTGTFYRN 534
 DB 462 DSGTTLSTTAGSITITNLGINVDSLGLKQPVSLTAKGASNKVIVSGKLANLIDIEGNIYS 521
 QY 535 HSLRNPOSVDILELKASGTVTS-----TAVTPDPINGEKFKHYGQGTWGPVWGTSATTT- 589
 DB 522 HMFSDHQLFSLKITTVDADVDTNVDISSLIPVPAEDPENSEYGFQGWNV-VNMTTDTATNT 580
 QY 590 --ATFNWTKTGYIPNPERIGSLVPSNLWNAFIDISSLHVLMTANEGLOGDRAFWCAGLS 647
 DB 581 KEATATWTKGVFPSPERKSAVLCNTLWGVFTDIRSLQOLVEIGATGMEHKGQFWSSMT 640
 QY 648 NFFHKDSTKRRGFRHLGGYVIGNLTCSDKILSAAPCOLFGRDRDYFVAKNQGTVYG 707
 DB 641 NFLHKTGDNRRGFRHTSGGYVIGSAHTPKDDLTFACHLFARDKOCFIAHNSRTYG 700
 QY 708 GTLYYQHNET-----YISL-PCKLRPCSLSYVTPITPVLFSGNLSYTHTDNDLTKYTTY 761
 DB 701 GTLFKFKHSTLQPNYLRGLGRAKFSESAIEKEFRIPLALDVQVSFSHSDNRMETHYTSL 760
 QY 762 PTVKSGWGNDSFALEFGGRAPICL-DESALPQYMPFMKLQFYVAHQEGFKQGTREARFP 820
 DB 761 PESEGSWSNECIAGGIGLDLPVLNPNPLFTFIPQMKVEMVYVSQNSFFSSSDGRGF 820
 QY 821 GSSRLVNALPIGIRFDKESDQDA-TYNLTILGYTVDLVRSNPDCTTLTIRISGDSWKTFG 879
 DB 821 SIGRLNLISIPVGAKF-VQGDIGDSYVDLSGFFVSDVYRNNPQSTATLWMSPDWKIRG 879
 QY 880 TNLARQALVLRAGNHFCSNFEAFSQSFELRGSSRYNVVDLGAKYQF 928
 DB 880 GNLSRQAFLLRGSNNYVYNSNCELFGHYAMELRGSSRYNVVDVGTGLRF 928

RESULT 6

PMP2_CHLPN
 ID PMP2_CHLPN STANDARD; PRT; 841 AA.
 AC Q9Z3A1; Q9RB73;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable outer membrane protein pmp2 precursor (Polymorphic membrane
 DE protein 2) (Outer membrane protein pmp2 precursor (Polymorphic membrane
 DE PMP2 OR OMP7 OR CPN0013 OR CP0761).
 GN Chlamydia pneumoniae (Chlamydia pneumoniae).
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=83558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VR1310;
 RX MEDLINE=20007584; PubMed=10539856;
 RA Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
 RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;
 RT "Molecular biology of Chlamydia pneumoniae surface proteins and their

Db 1183 LTFKHESAKPPSALLTLTGAVDAVDHDPHCITSL-TNGTWSFTATNLSRQAFPAEASG 1241

Qy 894 HFCNSNFARSQSFELRGSRNNVNDLGAQYQF 928

Db 1242 HLKLLHGLDCFASGSCELSSRSRSTNACGTRYSF 1276

RESULT 8

PMPL_CHLPN STANDARD; PRT; 922 AA.

ID PMPL_CHLPN Q924H9; 922 AA.

AC Q924H9; Q924H9; 922 AA.

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE Probable outer membrane protein pmpl precursor (Polymorphic membrane protein 1) (Outer membrane protein pmpl precursor (Polymorphic membrane protein 1) OR OMP6 OR CPN0005 OR CP0770).

OS Chlamydia pneumoniae (Chlamydia pneumoniae).

OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.

OX NCBI_TaxID=83558;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=VR1310;

RA Christian G., Boesen T., Hjerno K., Dagaard L., Mygind P., Madsen A.S., Knudsen K., Falk E., Birkelund S.; "Molecular biology of Chlamydia pneumoniae surface proteins and their role in immunopathogenicity"; Am. Heart J. 138:S491-S495(1999).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CWL029;

RA MEDLINE=9206606; PubMed=10192388;

RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W., Olinger L., Grimwood J., Davis R.W., Stephens R.S.; "Comparative genomes of Chlamydia pneumoniae and C. trachomatis."; Nat. Genet. 21:385-389(1999).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=AR39;

RA MEDLINE=20150255; PubMed=10684935;

RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., White O., Hickey E.K., Peterson J., Utterback T., Berry K., Baes S., Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R., Gwin M., Nelson W., DeBoy R., Kolony J., McClarty G., Salzberg S.L., Eisen J., Fraser C.M.; "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39."; Nucleic Acids Res. 28:1397-1406(2000).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=J138;

RA MEDLINE=20330349; PubMed=10871362;

RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K., Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.; "Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CWL029 from USA."; Nucleic Acids Res. 28:2311-2314(2000).

CC -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES) (POTENTIAL).

CC -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AJ133035; CAB37075.1; -

DR EMBL; AE001585; AAD18163.1; -

DR EMBL; AE002237; AAF38570.1; -

DR EMBL; AP002545; BAA98215.1; -

DR PIR; E72131; E72131.

DR PIR; E86491; E86491.

DR TIGR; CP0770; -

DR InterPro; IPR003368; Chlamydia_PMP.

DR Pfam; PF02415; DUF145; 1.

DR TIGRFAMs; TIGR01376; POMP repeat; 5.

KW Outer membrane; Signal; Multigene family; Complete proteome.

FT SIGNAL 1 26 POTENTIAL.

FT CHAIN 27 922 PROBABLE OUTER MEMBRANE PROTEIN PMP1.

FT CONFLICT 14 14 F -> L (IN REF. 1).

FT CONFLICT 375 375 Y -> C (IN REF. 1).

FT CONFLICT 606 606 D -> N (IN REF. 1).

SQ SEQUENCE 922 AA; DFF2ABG333AB031C CRC64;

Query Match 30.2%; Score 1442.5; DB 1; Length 922;

Best Local Similarity 36.9%; Pred. No. 3.6e-71;

Matches 351; Conservative 156; Mismatches 381; Indels 63; Gaps 26;

Qy 5 FP-KFVSTFAIFPLSMIAETVLDSSASPCNGKNGFVSRESQEDAGTTVLFKGNVTLE 63

Db 8 FPLVFSFTLLSVFDTLSATISLTPEDSPHGDSQN--AERSYVQAGDVYSLTGDVVIS 65

Qy 64 NIPGTGTAITKSCFNNTKGLDTFTGNGNSLLFQTVDAGTVAGAAVNSVVDKSTT-PTGF 122

Db 66 NV--DNSALNKACFNVTSGVTFAGNHGGLYFNMISSGTYKEGAVLCCQDPQATRSGF 123

Qy 123 SLSFIASPGSSITTKGAVSC--STGSLSTKNVSLFKNFSTNDGGAITAKTSLTG 180

Db 124 STLSFIQSPGDIKEQG-----CLYSKVALMLNNVVRFEQNSKTKGAGISANVTIVG 178

Qy 181 TTMSALPSENTSSKKGAIQTSALTITGNGEVSFSDNTSSD--SGAAIFTEASVTTSNN 239

Db 179 NYDSVSPYQNAAT-FGGAIHSSGFLQAVNQAEIRPAQNTAKNGSGGALYSDGIDIDQN 237

Qy 240 AKVSFIDNKVTGASSSTTGMDSGCAICAYKTSTDTK---VTLGNQMLFSSNNTSITAG 295

Db 238 AYVLFRENE-----ALTATGKGAGVCCLTSGSGSTPVPVITFSDNQVLFERNHSINGG 292

Qy 296 GAIYVKLELASGGLTLFNRNSVNGTAPKGAIAIEDSGELSLSDSGDVFVLTGNTVTS 355

Db 293 GAIYARKLSSGGPTLFINNISYANQNLGAIAIDTGGEISLSAEGKITTFQGN--RT 350

Qy 356 TTPGTNRSSIDLGSATKALRSAGRAIYFDYDITGSSSTTVDVLKVNTPADSAQY 415

Db 351 SLPLFN--GIHLLQNAKFLKQAEANGYSIEFYDIT--SEADGSTQLNINIGDPKNK--EY 404

Qy 416 TGNITFTGEKLSATEADSKNLTKLLQPVTLTSGGTLSLKHGVTLLQTAQQAQADSLM 475

Db 405 TGTILFSGEK---SLANDPRDFKSTIPQNVNLSAGYLVIKEGAETVSKFTQSPGSHLV 461

Qy 476 DVGTTL-EPADTSTNNLVINISSIDGAKKAK-IETKATSKNLTLSTGTTILLDPTGTYE 533

Db 462 DLGTLKLIASKEDIAITGLAIDIDSLSSSTAAVIAKANTANKQISVTDSEIISITGNAYE 521

Qy 534 NLSLRNPQSYDILELK--ASGTVTSTA-----VTPDPIMGEKHYGYQGTWGPVW-GTG 585

Db 522 DLRRNSQTPEPLLSLEFGAGGSVTVTAGDFLPVSP-----HYGFGQNW-KLAWTGTG 572

Qy 586 ASTTATFNWTKGYIPNPERIGSLVPNSLWNAFTDISSLHYLMETANEGLQDRAPWCAG 645

Db 573 -NKVGEFFWKINKYPRPEKEGNLVPNLLMGNADVRLMQVQVTHASSLQTDRLGWLIDG 631

Qy 646 LSNFPHKDSKTRGRFHLSSGYYVIGNLHTCSDKILSAAPQLFGDRDRDYFAKNGQTV 705

Db 632 IGNEFFHVSASEDNIRYHNSGGYVLSYNNETPKHYTSMASFSLFSRDKQYAVSNNEVRM 691

Qy 706 YGGTLYTOHNET-----YISLPCKLRPCSLSYVETIP-VLPSGNI-SYTHTDNDLKTKY 758

Db 692 YLGSYLYQYTTSLGNIIFRYASRNPNVNVGILSRRLQNPLMIFHLCAYGATNDKMTDY 751

Qy 759 TTYPTVKSGWNSDFALEFGGRAPICLDSEA-LFEQYMPMKLOFVYAHQBGFKQGTGA 817

```

Db 752 ANFPWVKNWRNNCAIECGSMPLLVFENGRLFOGAIPFMKLQVLVAYQGFKEITADG 811
QY 818 REFGSSFLNALPIGIRPKESDCQADATNLILGYTVDLVRNPDCITTLIRISGSWKT 877
Db 812 RRSNGSLTISVPLGRFRFKLALSQVLYDFSFYIPDFRKPSCAAALVSGDSWL 871
QY 878 FGTNLQALVLRAGNHFCNSNFAPQSFFELRGSSRNYNVDLGAQYOF 928
Db 872 PAAHVSRAHFGVSGTGYHENDYTELLCRGSIETCRPHARNYINCSKFRF 922

RESULT 9
PM13 CHLPN
ID PM13 CHLPN STANDARD; PRT; 973 AA.
AC Q92896; Q9K2A1; Q92410;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable outer membrane protein pmp13 precursor (polymorphic membrane
DE protein 13) (Outer membrane protein 14).
GN PMP13 OR OMP14 OR CPN0453 OR CP0299.
OS Chlamydia pneumoniae (Chlamydiophila pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiaceae; Chlamydiaceae; Chlamydiophila.
OX NCBI_TaxID=83558;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marache R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [4]
SEQUENCE OF 1-262 FROM N.A.
RC STRAIN=VR1310;
RX MEDLINE=20007584; PubMed=10539856;
RA Christiansen G., Boesen T., Hjerno K., Dagaard L., Mygind P.,
RA Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT "Molecular biology of Chlamydia pneumoniae surface proteins and their
RT role in immunopathogenicity.";
RL Am. Heart J. 138:S491-S495(1999).
CC -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL; AB001629; AAD18595.1; -
DR EMBL; AB002191; AAP38156.1; ALT_INIT.
DR EMBL; AF002546; BAA98660.1; -
DR EMBL; AJ133034; CAB37074.1; -
DR PIR; B86547; B86547.
DR PIR; F72076; F72076.
DR PHCI-2DPAGE; Q92896; -.
DR TIGR; CP0299; -.
DR InterPro; IPR005546; Autotransporter.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter_1.
DR Pfam; PF02415; DUF145; 3.
DR TIGRFAMs; TIGR01376; POMP repeat; 7.
KW Outer membrane; Signal; Multigene family; Complete proteome.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 973 PROBABLE OUTER MEMBRANE PROTEIN PMP13.
FT CONFLICT 258 258 N -> Y (IN REF. 4).
SQ SEQUENCE 973 AA; 102761 MW; E02A69F611DEBFE2 CRC64;

Query Match 28.9%; Score 1377.5; DB 1; Length 973;
Best Local Similarity 34.8%; Pred. No. 1.3e-67;
Matches 350; Conservative 153; Mismatches 391; Indels 113; Gaps 25;

QY 1 MKSSPKKVFST-----FAIFPLSMIAETVLDSSASFGCKNGKNGFVRESQEDAGTYVL 55
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 56
1 MKTIRKFLISTTLAPCFA---STAFTVEVIMPSNFENFDGSSGKIFPYTTLSDPRGTLCI
QY 56 FKGNTVLENIPGTGTAITKSCFNNTKGDLTFTFGNGSLLFQTVDACTVAGAAVNSVWDK 115
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 116
57 FSGDLVIANLDNAISRTSSCSFNAGALQILKGGVFSFLNI-RSSADGAISVITQN 115
QY 116-----STTFPGSLSF-----IASPGSITTGKAVSCSTGSLSLTKNVSLFKNFS 164
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 174
116 PELCPLSFGSPQMIFDNCELSLTSASNVIPHASAIYATTPML-FTNNDILFYQNS
QY 165 TDNGGAITAKTILSLGTTMSALFSENTSSKKGATQTSALTITGNQGVSFSD----- 218
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 219
175 AGFGAARTSITIENTKSLFNGSGISNGGALTGSAAILNNSAPVIFSTNATGY 234
QY 219-----NTSSDSGAAIFTEASVTISNAKVSFIDNKVKGASS- 254
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 294
235 GGAIVLTGSSMLTSGNLGVLVFNSSRGGAIYANGVTFNNSDLTFQNTASPNLSL 294
QY 255-----STTGDMG-GGAI-C---AYKSTDTKVLTGQMLLFNSNTSTTAGGAIY 299
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 354
295 PAFTPTPTTAVTPLLGYGAIFCTPPATPPPTGVSITGSESVTFLENIASEQGALY 354
QY 300 VKKLKELASGLTLFSESNVNGGTAPKGAIAIEDSGELSLSDSGDIYVPLGN-TVTSITP 358
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 409
355 GKKSIDSNKSTIFL-----GNTAGKGAIAIPESGELSLSANQGDILFNKNLSITSGTP 409
QY 359 GTNRSSIDLGTSAKMTALRSAGRAIYFYDPTTGG--SSTVTVDVLKN-ETPADSALQY 415
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 465
410 --TRNSIHFGDKAFATLGGATGYLYFYDPTSDLSAASAAATVNVNPRASADCA-Y 465
QY 416 TGNIIPTGEKLSEREAASKNLTKLQPVLSGGTSLSKHGVTLQTAFTQADSRLEM 475
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 525
466 SGTIYVSGETLTATEAATPANATSTLNQKLEGGTALRNCGATLVNHFNTQDEKSVVIM 525
QY 476 DVGTITLEPADTS-----TNNLVINISSIDGAKAKIETKATSKNLTSGTITLLDPT 528
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 585
526 DAGTTLATNGANNVDGAILNKLVINDSLDGTAKAAVVNVQSTNGALTISGTLGVKNS 585
QY 529 GTTFVNHSLRNP--QSYDILELKA-SGTVTSTAVTPDPIMGEKHYGQGTGPIVMCTG 585
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 644
586 QDCDCHGMFKNDLQOVFIIELEKATSNVTITDLSGNTNGYQQSPYGGQTFWEFTIDTT- 644
QY 586 ASTTATFNWTKTYGIPNPERIGSLVPSNLWNAFIDISSHYLMETANEGLOQDRAFWCAG 645
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 702
645 -THTVTGNWKKTYGLPHEPRLAPLIPNSLMANVIDLRAVSQAASADGEDVFG-KQLSITG 702
QY 646 LSNPHKDKSTKTRGRFHLSCGYVIGGNLHTCSDKILSAAFCQLFGRDRDYFVAKNQGTV 705

```

```

Db 703 ITNFFHANTGDARSYRHMGGVLYINTYTRITPDAAALSLGFLTKSKDYLVGHCHSNV 762
QY 706 YGGLTYXQHNETYISLPCKLRPCSLSVVPEIPVLESGNLS---YTHTDNDLTKYTYTP 762
Db 763 YFATVYSNITKSLFG-----SSRFSGGTSTRTVYSRNEKVKTSYTKLP 806
QY 763 TVAGSGMNSDFALEFGGRAPICLDESAL-PEQYMPFMKLQFYVAHQEGFKQGTAREBG 821
Db 807 KGRCSMNSNCWLGELEGNLPITLSSRLNLKQIIPVKAEBAYATHGGIOENTPEGRIFG 866
QY 822 SSRVNLALPIGRFKESDCQATYNLIGYTVDLVRNPDCTTTLRISGDSWKTFCGN 881
Db 867 HGHLNVAVPVGRFGKNSHNRPDFYIIIVAYAPDVYRHNPDCTTLPINGATWTSIGNN 926
QY 882 LARQALVLRAGNHFCFNSNPEAFSQSFELRGSSRNYNVDLGAQYQF 928
Db 927 LTRSTLLVQASSHTSVNDVLEIFGHCGCDIRRTSRQYTLIDIGSKLRF 973

```

RESULT 10

```

PMPG CHLTR
ID PMPG CHLTR STANDARD; PRT; 1013 AA.
AC O84879;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Probable outer membrane protein pmpG precursor (Polymorphic membrane
DE protein G).
DE PMPG OR CT871.
GN Chlamydia trachomatis.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/Cx;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RA "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis."
RL Science 282:754-759 (1998).
CC -1- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE001360; AAC68469.1;
CC PIR; G71460; G71460.
CC PHCI-2DPAGE; O84879;
CC InterPro; IPR006315; Autotransport.
CC InterPro; IPR005546; Autotransporter.
CC InterPro; IPR003368; Chlamydia_PMP.
CC Pfam; PF03797; Autotransporter; 1.
CC Pfam; PF02415; DUF145; 1.
CC TIGRFAMs; TIGR01414; autotrans barl; 1.
CC TIGRFAMs; TIGR01376; POMP repeat; 6.
KW Outer membrane family; Complete proteome.
FT SIGNAL 1 27
FT CHAIN 28 1013 PROBABLE OUTER MEMBRANE PROTEIN PMPG.
SQ SEQUENCE 1013 AA; 107366 MW; F0927743COA651DD CRC64;

```

Query Match 23.5%; Score 1120.5; DB 1; Length 1013;
 Best Local Similarity 29.6%; Pred. No. 1.4e-53;
 Matches 313; Conservative 160; Mismatches 409; Indels 177; Gaps 33;

```

QY 1 MKSFPKPFSTFAIFPLSMIA-----TETVLDSSASFDGNK-NGNFSVRESQEDAGTYL 55
Db 1 MOTSFHKKFFLWILAYSCCSLGGGYAAEIMIPQGIYDGETLTVSPYTVIGPDSGTVF 60
QY 56 PKGNVTLENIPGTGTAITKSCFNNTKGDLLFTPTGNSNLLPQTVDAGTVAGAAVNSVVDK 115
Db 61 SAGELTLTKNDNSIAALPLSCFGLNLLGSFTVLGRGHSALTFFENIRTS-NGAALSDSANG 119
QY 116 STTIFIGSLSLF-----IAGPGSSITTKGAVSCSTGSLSTKNVSL 157
Db 120 LFTTIEGKELSFNCNLSLLAVLPAATTNNGSQPTTSTPSNGTIYSKT-DLLLLNNEKF 178
QY 158 LFSNFTDONGAITAKTSLTGTMTSALFSENSTSKKGAIQTSOALITITNGQEVSP- 216
Db 179 SFYNLVSVDGGAIDAKSLTVQGISKLCPQENTAQADGACQVTVSFSAMANEAPIAFI 238
QY 217 -----SDNTSSDS-----GAAIFTEASVT 235
Db 239 ANVAGVRGGGIAVQDQGGQVSGSTSTEDPVVFSRNTAVEPQGNVARVGGGIYSYGNVA 298
QY 236 ISNNAKVSFIDN-----KVTGASSSTTGDM--SGGAI-C-----AYKSTDTTKVT 277
Db 299 FLNNGKTLFLNNVASPVVIAAEQPTNGQASNTSDNYGDGGAIFCKNGAQAGSNNSSVS 358
QY 278 LTGNQMLLFSNNTSTTAGGAIYVKLELASGLTLFSRNSVNGGTAPKGAIAIEDSGEL 337
Db 359 FDGEGVFFSNAAGKGGAIYAKKLSVANCGPVQVFLGNIAN-----DGAIIYLGESGEL 413
QY 338 SLSADSGDIVPLGN---TVTSTTTPGTIN-----RSSIDLGTSAKMTALRSAGRAIFYDP 389
Db 414 SLSADYGDIIIPDGNLKRKTAKENADVNGVTVSSQAIISMGSGGKITTLRAKAGHQLFNDP 473
QY 390 ITTGSSTT---VTDVLKVNETPADSALOYTGNIIFTGEKLSSETAADSKNLSKLLQPV 445
Db 474 IEMANGNNQPAQSSEPLKINDGEG---YTGDIVFA-----NGNSTLYQNV 515
QY 446 TLSGGLSLKHGVTLOQAFQQAQDSRLMDMDVGTLE-----PADTS---TINNVL 493
Db 516 TIEQRIVIREKAKLSVNSUSQTGS-LYMEAGSTLDFVTPPQPQPPAANQLITLSNLH 574
QY 494 INISSI---DGA-----KKAKIETKATSKNLTSLGTTIPLLDPTGTFYENHS-LR 538
Db 575 LSLSLLANNAVTPNPPAPDQSHPAIGSTTAGSVTISGPIFFEDLDDTAYDRYDWLG 634
QY 539 NPQSDVILELKASCTVSTAVTPDPIMGEKF-HYGYOGTWGPVIMGTGASTATF----N 593
Db 635 SNQKIDVLKQL-GTQPSANAPSDLTIGNEMPKYGYGGSW-KIAWDPNNTANGPYTLKAT 692
QY 594 WTKTGYPNPERIGSLVPSNLWNAFIDITSLHYMETANEGLOQDRAPWCAGLSNFFHKD 653
Db 693 WTKTGYNFGPERVASLVPSNLWMSILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYHD 752
QY 654 STKTRGRPHLSGGVVGUNLHTCSDKLSAFCQLFGRDRDYPVAKNQGTVTGTYLQY 713
Db 753 RDALGGYRYISGGYSLGANSYFGS-SMPGLAFTEVFGRSKDYVVCNSHHACISGYV-- 809
QY 714 HNETVISLPCKLPCSLSVVPEIPVLFSG---NLSYTHTDNDLTKYTYTPYTVKSGWN 770
Db 810 -----LSTQALCG-SY-----LFGDAFIRASYFGNQHMKTSYTTAEBSDVWRDN 854
QY 771 DSFALEFGGRAPICLDESALF-EQYMPFMKLQFYVAHQEGFKQGTAREFGSSRLYNLA 829
Db 855 NCLVGEIGVGLPIVITPSKLYLNEIRPFVQAEFSYADHESFTBEGDQARAFSRGHLNLS 914
QY 830 LPIGRFKESDCQATYNLIGYTVDLVRNPDCTTTLRISGDSWKTFCGNLARQALVL 889
Db 915 VEVGKVPDRCSSTHPNKYSFMGAYICDAYRTISGTQTLTLLSHQETWTTFDAPFLARHGVI 974
QY 890 RAGNHFCFNSNPEAFSQSFELRGSSRNYNVDLGAQYQF 928
Db 975 RGSWYASLTSLNIEVYHGRYRDTSRGYGLSAGSKVRP 1013

```



```

RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,
RA Linner K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39."
RT Nucleic Acids Res. 28,1137-1406(2000).
CC -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE002294; AAF39136.1; -.
CC PIR: F81721; F81721.
CC TIGR: TCO267; -.
CC InterPro; IPR003368; Chlamydia_PMP.
CC Pfam; PF02415; DUF145; 1.
CC TIGRFAMs; TIGR01376; POMP repeat; 6.
KW Outer membrane; Signal; Multigene family; Complete proteome.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 867 PROBABLE OUTER MEMBRANE PROTEIN PMPI.
SQ SEQUENCE 867 AA; 95017 MW; 557994185A9E5652 CRC64;

Query Match 19.1%; Score 911; DB 1; Length 867;
Best Local Similarity 29.5%; Pred. No. 2.7e-42;
Matches 281; Conservative 144; Mismatches 384; Indels 144; Gaps 30;

QY 17 PLSMIATETVLDSSAFDGNKNG-----NFS---VRESQED-----AGTYILF 56
DB 18 PTAIFGQDALOKSALITKPNPSIVCTFLEDCMTENFSPALLSHARQDDPLYIINTH-- 75
QY 57 KGNVTLENT-PGGTGAITKSCFNNTKGDLTFTGNGSNLLFQTVDAGTAVAGAAVNSVVDK 115
DB 76 --NWFVSNLHPST-----NEERFLKEKGLSI-----Q 101
QY 116 STTFIGFSSLSFTASPGSSITTCGKAVSCSTGSLTKVNSLLPKNFSTNGAITAKT 175
DB 102 DFRFLFTDCSSSTEDSPSILYHK-----NGQLFLRNNGNMFYRNHSEGGGALSTDA 155
QY 176 LSLTGTMTSALFSENTSSKKGAIQTSDALTTGNQGEVSFSDNTSSDSGAIFTEASVT 235
DB 156 LFLQHYLYFTNFENSARKNGAIQ-AQLTSLSRNVSSLSFBRNANLNGAICCNQLIC 214
QY 236 ISNNAKVSPFDNKVTGASSSTTGDMSGGAICAYKTSITDTK---VLTGQMQLLFSNNTST 292
DB 215 SGNVNPLFFTTNSA-----LNGGAICCNQNEKGLSLAYNQETLFGNSAK 264
QY 293 TAGAIVVKKLELASGLTLFSPNSVNGGTAPKGGAIATEDSGELSLSDSGDVIPLGNT 352
DB 265 EKGGAIVYTHVLRHNGVPFVNNS-----AKLGGAIATQSGGSLSIAGGGSVLFQNNNS 319
QY 353 VTSTTPTGTRNSIDLGTSAWTAALRSAGRAIYFYDPIITG-----SSTVTVDVLKYN 406
DB 320 CHFSDQGTVENAIVLEKNALLSLEARHG-DILFFDPIQEVVSPFSTTSALTPLRI-Q 377
QY 407 TPADSAIQYTGNIIFTGKELSETEAADSKNLTSLKLLQPVTLGGTSLKHGVTLQTPAT 466
DB 378 TTNTRA-----VIFSSENLSKEKTEA-NLSIKIQPIELQSGCIVLKDRIVLSAPSLS 430
QY 467 QQADSRLEMDVGTTLBPA DSTNNIVINISSIDGAKAKIEFKATS-KNLTLSGTITLL 525
DB 431 QAFQALLVMDVGLSTLTSSDLKUTLSIPHSITDENSIVSGIPTSLQIKFLSNS----- 486
QY 526 DPTGTGFYENHSLRNPOSYDILELKASQTVTSTAVTPD-FIMGKFKHYGQGTWGPVWGT 584

```

RESULT 13

```

PMPI_CHLTR
ID PMPI_CHLTR STANDARD; PRT; 878 AA.
AC 084882;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable outer membrane protein pmpi precursor (Polymorphic membrane
DE protein 1).
GN PMPI OR CT874.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=D/UM-3/Cx;
RX MEDLINE=99008089; PubMed=9784136;
RA Stephens R.S., Kalkan S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis."
RL Science 282:754-759(1998).
CC -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE001361; AAC68472.1; -.
CC PIR: B71460; B71460.
CC InterPro; IPR003368; Chlamydia_PMP.
CC Pfam; PF02415; DUF145; 1.
CC TIGRFAMs; TIGR01376; POMP repeat; 6.
KW Outer membrane; Signal; Multigene family; Complete proteome.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 878 PROBABLE OUTER MEMBRANE PROTEIN PMPI.
SQ SEQUENCE 878 AA; 95592 MW; DF1FIA3170EE48B CRC64;

```

```
Query Match      18.1%; Score 863; DB 1; Length 878;
Best Local Similarity 30.1%; Pred. No. 1.1e-39; Indels 112; Gaps 23;
Matches 260; Conservative 124; Mismatches 368;

Qy 118 TFIGFSSLPFASPGSSITTKGAVS-----CST-----GSLSTTKN 154
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 74 SYCWFSKLIHTDPKEALFKKGLDLSIQNFRFLSFTDCSSKSSPSIIHKQKQLSLRN 133

Qy 155 VSLFLSKNFTDNGGATATKLSLTGTMTALSFNTSSKKGAIQTSDALTITNGGEV 214
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 134 GMSFCRNHAEGGGAISADAFSLQHNLYLFTAFENSSKNGGAIG-AQTFSLSRVSP 192

Qy 215 SFSDNTSSDGAIFTEASVTISNAKVSFIDNKVTGASSSTTGDMGGGAICA---YKTS 271
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 193 SFARNADNGGALCCSNLCSGNVNPFTFGNSAT-----NGGAICISDLNLS 242

Qy 272 TDKVTLTGQMLFNSNNTSTTAGGAIYVKKELASGGLTLFRNSVNGGTAPKGAIAI 331
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 243 EKGSLSLACNQETLFASSAKERKGAIAKHMVLRVNGPVSVFINNS-----AKIGGAIAI 297

Qy 332 EDGSELSLSDGDIPLGNVTSTTPTNRSSIDILGTSKMTALRSAGRAIFYDPDPT 391
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 298 QSGSLSILAGEGVLFQNNQSORTSQGLVNAIYLEKDAILSLSEARNG-DILFPDPIV 356

Qy 392 TGSSTTVDVLKVNTPADSAALQVT-----GNIIFTGEKLSATEAA 432
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 357 QESSS-----KESPLPSSLOASVTSPTPATAPLVIOTSANRSVFSSELSSEK 408

Qy 433 DSKNLSKLLQPTVLTGGTSLKHGVTLOQTAAQDSRLMDVGTTLPEADPTSTNNL 492
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 409 PD-NLTSQLOQPIELXSGRLVLDRAVLSAPLSQDPQALLIMEAGTSLKTSDDLKATL 467

Qy 493 VINISSIDGAKKAKIETKATSKNLTLSGTTLLDPTGFVENHSLNPQSYDILELKASG 552
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 468 SIPHLSDTEKSVTH-----APLSIQKIFLSNGDENFYENVVELLSKEQNNIPLTLS- 522

Qy 553 TVTSTAVTPDIPMERKHYGQGTWGPVW---GTGASTTATFNWTKGTGYIPNPERIGSL 609
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 523 KEQSHLHLPDGNLSS--HFQYQGDW-TFSWKDSDEGHSLIA--NWTPKYVVPHERQSTL 577

Qy 610 VPSNLWAFIDISLHYLMETANEGLQGRAP--WCAGLSNFF--HKDSTKTRRGFRHLS 665
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 578 VANTLWNTSDMQAVQSMINTIAHG--GAYLFTGWSAVSNLIFYAHDSSGKPIDNWHRS 635

Qy 666 GGVIIGNLHTCSDKILSAFQCLFGRDRDYFAKNQGTGYGGLTYQHNETYISLPCKL 725
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 636 LGVFLGISTHSDHDFCLAGQLLKSDSFTSTETTSYIATVQAQ-----LATPL-- 688

Qy 726 RPSLSYVTEIPVLFSGNLSYTHTDNDLTKYTYPTVK--GSWGNDSFALEPGGRAPIC 784
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 689 -----MKISAQACYNESITHELTKYRFSKSGFSGSWHSVAVSVEVCASIPIV 735

Qy 785 LDESALFEQVMPMKLQFYVAHQEGKQGTAREFGSGRLVNLALPIGIRFDKESDCOD 844
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 736 SNGSGLFSSFSIFSKLQGFSGTQGFESSGSEIRSFSSFRNISLPMGITTEKKSQKTR 795

Qy 845 ATYNLTGLTVDLVRNPDCTTLTIRSGDSWKTFGTNLARQALVLRAGNHFCFNSFEAP 904
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 796 NYIFYLGAIQDLKRVEGSPVLLKNVSWDAPMANLDSRAMVFLTLNORALH-RLQTL 854

Qy 905 SQFSFELGSSRNYNVDLGAKYQF 928
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 855 LNVSYVLRGQSHSYSLDLGTYRF 878
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 14

```
PM21-CHLPN
ID PM21 CHLPN STANDARD; PRT; 1609 AA.
AC Q9Z6U5; Q9RB58;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable outer membrane protein pmp21 precursor (Polymorphic membrane
```

```
protein 21).
PMP21 OR CPN0963 OR CF0897.
GN Chlamydia pneumoniae (Chlamydophila pneumoniae).
OS Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OC NCBI_TaxID=83558;
RN [1] SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
OLinger L., Grimwood J., Davis R.W., Stephens R.S., Bowman C., Dodson R.,
RA "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Uchterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Bisen J., Fraser C.M.;
RA "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shitai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RA "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
(POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; AE001676; AAD19099.1; -.
DR EMBL; AE002248; AAF38684.1; -.
DR EMBL; AF002548; BAA99171.1; -.
DR PIR; A86611; A86611.
DR PIR; H72013; H72013.
DR PHCI-2DPAGE; Q9RB58; -.
DR TIGR; CP0897; -.
DR InterPro; IPR006315; Autotransport.
DR InterPro; IPR005546; Autotransporter.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF02415; DUF145; 1.
DR TIGRFAMs; TIGR01414; autotrans bar1; 1.
DR TIGRFAMs; TIGR01376; POMP repeat; 13.
KW Outer membrane; Signal; Multigene family; Complete proteome.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 1609 PROBABLE OUTER MEMBRANE PROTEIN PMP21.
FT CONFLICT 420 420 I -> M (IN REF. 3).
SQ SEQUENCE 1609 AA; 170865 MW; 2604C3E9FC4024CB CRC64;
```

```
Query Match      14.4%; Score 685.5; DB 1; Length 1609;
Best Local Similarity 24.8%; Pred. No. 1.2e-29;
Matches 254; Conservative 144; Mismatches 372; Indels 253; Gaps 36;

Qy 60 VTLENIPGTGTAITSKCF--NNTKGLTFTGNGSLLFTQVDAGTVAGAAVNSVVDKST 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 686 VSSDIRGGGAILAQHIFITNT-GNLRFSGN-----LGGGESSTV---- 726
```

[illegible]

```

AC QZ2895; Q9RB63;
AD 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable outer membrane protein pmpl4 precursor (Polymorphic membrane
DN protein 14).
GN PWE14 OR CPN0454 OR CP0298
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OC NCBI_TaxID=83556;
[1]
FN SEQUENCE FROM N.A.
RP STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kalman S., Mitchell W., Warathe R., Lammel C., Fan J., Hyman R.W.,
FA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
[2]
FN SEQUENCE FROM N.A.
RP STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brubham R.C., Shen C., Gill S.R., Heidelberg J.F.,
FA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
FA Gwin N.M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
[3]
FN SEQUENCE FROM N.A.
RP STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
FA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001629; AAF18596.1; -;
DR EMBL; AE001291; AAF38155.1; -;
DR EMBL; AP002546; BAA98661.1; -;
DR PIR; B81593; B81593.
DR PIR; C86547; C86547.
DR PIR; G72076; G72076.
DR PHCI-2DPAGE; Q92895; -.
DR TIGR; CP0298; -.
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF02415; DUF145.1.
DR TIGRFAMs; TIGR01376; POMP repeat; 6.
KW Outer membrane; Signal; Multigene family; Complete proteome.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 978 PROBABLE OUTER MEMBRANE PROTEIN PMP14.
FT CONFLICT 378 AA; 103654 MW; 6EEE1429999D3019 CRC64;
SQ SEQUENCE 978 AA; 103654 MW; 6EEE1429999D3019 CRC64;

```

Job time : 21 secs

Db 2 PLSPKSSPCLLACLCASCAFETRIGGNFVPITNQ--GEEILLTSDFVCNFIUG--- 56
QY 71 AITKSCFNNTKGDILFTNGNSLLFOTVADAGTVAGAAVNSVVDKSTTFIFGSSLSFIAS 130
Db 57 ASFSSSPINSNNLSLLGKGLSLFTSCQAPTSNYALLSAA--ETLTFKNFSSINFNGN 114
QY 131 PGS---SITTKGNVSCSTGSLSTKXVNSLLFSKNPSTDNGGAIKATLSTGTT---M 183
Db 115 QSTGLGLIYKGOIVFOISIKDLIFITNRYVSPASVTTSPAITVT--TCASALOPT 171
QY 184 SALFSENTS-----SKKGAIOTS DALTI--TCNGGEVSPSDNTSSDGAAIPTFA 232
Db 172 DSLTVENISQSIKFFGNLANFGSAISSPTAVVKPINNTATMSFHNFTSSGGVIYGS 231
QY 233 SVTISNNAK-VSFIDNKVTGA-----SSSTTGMMSGGAI CAYKTSDDTK----- 275
Db 232 SLLFENNNGGCIIFTANSCVNSLKGVPSSGYALGSGGAI CIPGTPELKNQKCTFSY 291
QY 276 -----VLTGNQ-MLLFSNNTSTTAGGAIYVKKLELASGGLTLPFRNSVNG 320
Db 292 NGTPNDAGAIYAETCNIVGNQGALLDSNTAARNGGAI CAKVLNIQGRGPIEFSRN--- 347
QY 321 GTAPKGGAI AI-----EDSGELSLSADSGDIVFLGNTVTSPTGTRNSSIDLGTSAK 372
Db 348 -RAEKGGAIFIGPSVGPAPKQTSLTITLASEGNIAFOGN-MLNTKPGI-RNAITVEAGGE 404
QY 373 MTAIRSAAGRAIYFDPITGSSSTVTDVLKVNETPADSALQYTGNIIFTGEKLSSETAA 432
Db 405 IVSLSAQGGSRLLVFYDPITHSLPTTSPS-----NKDITINANGASGVVFTSKGLSSTELL 460
QY 433 DSKNLTSLKLOPVTLSGGLTSLKHGVTLOQAFQOADSRLMDVGTTL-----EPAD 485
Db 461 LPANTTITLLGTWKIAGELKIDTNAVNVNLGFATQSGGQLTSGGGTGLGLATPTGAPAA 520
QY 486 TS-TINNVLVINISSIDGAKAKIETKATSKNLTLSGTTILLDPTGT-FYENHSLRNPOSY 543
Db 521 VDFITIGKLA DPFSFLKRDVFSASVNA GTKNVLTGALVLDHVDVTDLYDMVSLQSPVAI 580
QY 544 DILELKASGTVTSTAVTPDPIMGEKFKHYGQGW-----GP----- 579
Db 581 PIAVFKGA-TVTKTGF-PDGEIATPSHYGQKWSYTWSRPLIIPADPGGPGGPPSPSAN 638
QY 580 ---IVMGTGASTTATFNWTKTYIPNPERIGSLVPSNLW-----NAFIDISSIHYLMET 630
Db 639 TLAVVNSDTLVST-----YILDERYGEIVSNLSWISFLGNQAFSDI--LQDVILLI 689
QY 631 ANEGLQGDRAFCAGLNFPHKDKSTRGRHLSGGYVIGGNLHTCSDKILSAAFCQLF 690
Db 690 DHFGL-----SITAKALGAYVEHTPRQGEHGFSGRYGQYQAALSNNYTDHTTLGLSFGQLY 745
QY 691 GRDRDYFVAKNQGVYGGTLYYQHNETYIISLPCKLRPCSLSY-----VTEIPVLPSGN 744
Db 746 GKT-----NANPYDSRCSEQMYLLSFTGQFPITVQKSEALISWK 784
QY 745 LSYTHTDNDLTKY---TTPYTKVSGNDSPALFEGGRAP I---CLDESALFEQY--MP 796
Db 785 AAYGYSKXHLNTYLRPDKAPKQGGQWHNNYSYVLSAEHPFLNWCLLTRPLAQAWDLG 844
QY 797 FMKLQFYVHOGFKEGQTBAREFGSSRLVNLALPIGIR-----FDKESDCQDATYNLT 850
Db 845 FISAFLGGWQSKFTENGDLQRGSRGKGYNSVLPICSSQWFTPPKK-----APSTLTIK 900
QY 851 LGYTVDLVRNPDCITTLRISGDSWKTFCTNLARQALVLRAGNHFCNSNFEAFSPSFE 910
Db 901 LAYKPDYIRYNPHNIVTVWSNQESTSISGANLRHGLFVQIHDVVDLTDTOAFLNVTFD 960
QY 911 LRGSSRNYNVDLGAKYQF 928
Db 961 GKNGFTNHRVSTGLKSTF 978

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2003, 10:35:25 ; Search time 45 Seconds

(without alignments)
5321.618 Million cell updates/sec

Title: US-09-428-122-2

Perfect score: 4774

Sequence: 1 MKSSPFKVFSTFAIFPLSM.....FELGRSSRYNVDLCAKYQF 928

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp archaea:*
- 2: sp bacteria:*
- 3: sp fungi:*
- 4: sp human:*
- 5: sp invertebrate:*
- 6: sp mammal:*
- 7: sp mhc:*
- 8: sp organelle:*
- 9: sp phage:*
- 10: sp plant:*
- 11: sp rodent:*
- 12: sp virus:*
- 13: sp vertebrate:*
- 14: sp unclassified:*
- 15: sp virus:*
- 16: sp bacteriap:*
- 17: sp archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3991	83.6	772	16	Q9RB71
2	1915	40.1	926	2	P71135
3	1656.5	34.7	839	2	P77792
4	1641.5	34.4	847	2	P71132
5	1604.5	33.6	846	2	P71133
6	1237	25.9	602	2	Q8VU49
7	1219.5	25.5	601	2	Q8VLS7
8	1175	24.6	700	2	Q8VU50
9	1172.5	24.6	581	2	Q8VU48
10	1017.5	21.3	445	16	Q9RB67
11	935.5	19.6	649	2	P71134
12	916	19.2	359	16	Q9JSK6
13	821	17.2	494	16	Q9RB68
14	792	16.6	427	16	Q9RB70
15	768	16.1	186	16	Q9RB72
16	580	12.1	1530	2	Q93Q88

17	379.5	7.9	252	2	Q8VLH7
18	372	7.8	257	2	Q8VLF9
19	328	6.9	151	16	Q9JSK7
20	310.5	6.5	1213	16	Q9JH8
21	308.5	6.5	162	16	Q9RB69
22	307.5	6.4	1713	3	Q8TGE1
23	302.5	6.3	3242	2	Q8G9X9
24	300.5	6.3	2201	2	Q8GF46
25	284	5.9	2283	2	Q8VQ99
26	282.5	5.9	2232	5	Q8IFX6
27	282	5.9	1341	16	Q8UAU1
28	277.5	5.8	1001	3	Q85164
29	275	5.8	1275	5	Q76602
30	275	5.8	1557	2	Q9RNI2
31	275	5.8	2554	16	Q8YDM7
32	274	5.7	2893	16	Q25063
33	273	5.7	392	16	Q9Z881
34	271	5.7	3225	16	Q8PRM0
35	270	5.7	3420	16	Q8FUS1
36	267.5	5.6	1672	16	Q8Y366
37	267.5	5.6	3961	16	Q8P942
38	267	5.6	1008	16	Q8KCG9
39	265.5	5.6	744	3	Q8TFG9
40	265	5.6	3930	16	Q98820
41	262.5	5.5	347	16	Q9RB61
42	261	5.5	2275	16	Q8NUJ3
43	260	5.4	2271	16	Q99QV4
44	258.5	5.4	888	5	Q25336
45	256.5	5.4	1758	16	Q9JMS5

ALIGNMENTS

RESULT 1

Q9RB71	ID	Q9RB71	PRELIMINARY;	PRT;	772 AA.
AC	Q9RB71;				
DT	01-MAY-2000	(TREMBLrel. 13, Created)			
DT	01-OCT-2000	(TREMBLrel. 15, Last sequence update)			
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)			
DE	Pmp 3.				
DE	Pmp 3.				
GN	Pmp_3_2				
OS	Chlamydia pneumoniae (Chlamydophila pneumoniae).				
OC	Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.				
OX	NCBI_TaxID=83558;				
RN	[1]				
RC	SEQUENCE FROM N.A.				
RP	STRAIN=J138;				
RX	MEDLINE=20330349; PubMed=10871362;				
RA	Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K., Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;				
RT	"Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CWL029 from USA.";				
RL	Nucleic Acids Res. 28:2311-2314(2000).				
DR	EMBL; AP002545; BAA98226.1;				
DR	InterPro; IPR006315; Autotransport.				
DR	InterPro; IPR005546; Autotransporter.				
DR	Pfam; PF03797; Autotransporter; 1.				
DR	Pfam; PF02415; DUF145; 1.				
DR	TIGRFAMS; TIGR01414; autotrans_bar1; 1.				
DR	TIGRFAMS; TIGR01376; POMP repeat; 6.				
SQ	SEQUENCE 772 AA; 82931 MW; 484FC56D635801EB CRC64;				

Query Match 83.6%; Score 3991; DB 16; Length 772;

Best Local Similarity 99.9%; Pred. No. 3.9e-179;

Matches 771; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 157 LLFSKNPSTNGAIIAKTLTGLTWTMSALFSENTSSKKGAIQTSDALTITNGQEVSF 216

Db 1 MLFSKNPSTNGAIIAKTLTGLTWTMSALFSENTSSKKGAIQTSDALTITNGQEVSF 60


```

QY 57 KGNVTLENIPGTGTAITKSCFNNTKGDLTFTGNGNSLLPOTVDAGTVAGAAVNSVVDKS 116
Db 61 TGNICAYAGLDSGLSSCFTDTAGNLSPLNGYTLTLCFNTITQSSHPGCAIVSGTNTK 120
QY 117 TTFIGPSLFIASPGSITTTGAGVSCSGSLTKNVSLFSSKFNSTDNCGAIAKTIL 176
Db 121 LDLSGSLFSCAYCPGPA--TGGAIK-AVENTIIDKNSLSLVFHKNCSTGEGGAIOCKAS 177
QY 177 SLRGTTMSALFSENTSSKKGAIQTSDALTTITGNGEVSPSDNTSSDSGAATFEASVTI 236
Db 178 S-----SEALKI 185
QY 237 SNNAKVSFDNKVTGASSSTTGDMGGCAICAYKTSTDTKVTLTGNQMLLFSNNTSTAGG 296
Db 186 EN-----NONLVFAENSSSSSGG 203
QY 297 AIYVKLELASGGLTLFNRNSVNGTAPKGAIAIEDS-GEISLSDSGDGVFLGNTVTS 355
Db 204 AIYADKLTIVSGGFTLFSNNSV8-ASSPGKGAICIKDSGECSLTADLGDITFDGNKIILK 262
QY 356 T---TPGTRSSIDLGTSAKMTALRSAAAGRAIFYDPIITGSSSTTVTDVLKVNETPADSA 412
Db 263 TNGGSPVTRNSIDLSSGKFTKLNAGEFGIFYDPIITGGG8-----DELNINK---QDT 315
QY 413 LQYTGNIIFTEGKLSERADSKNLTSLLOPVTLSGGTSLKHGVTLOTAQFTQADSR 472
Db 316 VDTGKIVFSGERLSDEEKVAANLAKSDFKQPKIGSGSLILKDGVTLETKSPTQEGAT 375
QY 473 LEMDVGTLEP-----ADTSTINNLINISSIDCA---KKAKIETKATSKNLTSGTITL 524
Db 376 VVMDLGTTLQTPSSGGETITLNLNDINVASLGGGVAPDPKAVEATESKTIVIN-AVNL 434
QY 525 LDPGTFTYHNHSLRNQSDVIDELK--AGTIVTSTAV-----TPDPIMGEKPHYGQGTW 577
Db 435 VDDNGNAYEPILAAQOPPTAIEVRSGSGSITKPTTNLENITYPPT-----HYGYQGNW 488
QY 578 GPIVWGTGAST---TATFNWTKTYTPNPERIGSLVPSNLWNAFIDISSLHYLMETANEG 634
Db 489 -TWTWKQSSAQEKATLTLWEQTYGSPNPERQSLVPLTWGFSFDIRAIONLMDISVNG 547
QY 635 LQDRAFWCAGLSNFFPHKOSTKTRFRHLSGGYVIGGNLHTCSDKILSAAFCQLPGRDR 694
Db 548 ADVHRGFWSGLGNFLHKSQSDFKFRHNSAGYALGVYAQTPSEVDFVSAAFQCLEGKDK 607
QY 695 DFVANKQGTIVGTYGTYQHNETYIISLPCKLRPCSLSYPT-----EIPVLFSG 743
Db 608 DYLVSNSSTVYAGSIYYQH-----ISYWNVTWNTLLQNTLIGRAEAPLVNA 652
QY 744 NLSYTHTDNDLTKY-----TTYPTVKSGWGNDSFALEFGRAPICLDESALFEQYM 795
Db 653 QLAYCHASNMTKNTMTDTPAPKTTYBEEKGDMGNDGCFGEVFAKAPI-ETASLLFDMYS 711
QY 796 PFMKLOFVYAHQBFKEQGT--EAREFGSSRLVNLALPIGIRFQKESDQCDATVNLTLGYT 854
Db 712 PFVKQLQVLAHQDDFKENNSDQRYFESNLTNLSMPIGVKLEFKSHKDTASNLTLAYA 771
QY 855 VDLVRNPDCITTLRLISGDS--WKTFTGLARQALVLRAGNHFCFNSNFEAFSQFSELR 912
Db 772 PDIVRSNPDCITALLVSPTSVAVWVTKANNLARHAFILQAGNYLALTRNTELFSEQFSELR 831
QY 913 GSSRNYNVDLGAQYQF 928
Db 832 GSCRTNIDUGSKIQF 847

```

RESULT 5

P71133

ID P71133

AC P71133

DT 01-FEB-1997 (TrEMBLrel. 02, Created)

DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE POMP91B precursor.

```

OS Chlamydia abortus.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
NCBI_TaxID=8355;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=S26/3;
RC MEDLINE=96406378; PubMed=8810511;
RA Longbottom D., Russell M., Jones G.E., Laineon P.A., Herring A.J.;
RT "Identification of a multigene family coding for the 90 kDa proteins
of the ovine abortion subtype of Chlamydia psittaci.";
RL FEMS Microbiol. Lett. 142:277-281(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S26/3;
RC MEDLINE=98187897; PubMed=9529048;
RA Longbottom D., Russell M., Dunbar S.M., Jones G.E., Herring A.J.;
RT "Molecular cloning and characterization of the genes coding for the
highly immunogenic cluster of 90-kilodalton envelope proteins from the
Chlamydia psittaci subtype that causes abortion in sheep.";
RL Infect. Immun. 66:1317-1324(1998).
DR EMBL; U65943; AAC15923.1; -
DR InterPro: IPR003368; Chlamydia_PMP.
DR Pfam: PF02415; DUF145; 3.
DR TIGRFAM; TIGR01376; POMP_repeat; 3.
KW SIGNAL.
FT CHAIN 1 16 POTENTIAL.
FT SIGNAL 17 846 POMP91B.
SQ SEQUENCE 846 AA; 90834 MW; 4CDC31DC03C2964E CRC64;

Query Match 33.6%; Score 1604.5; DB 2; Length 846;
Best Local Similarity 38.7%; Pred. No. 2.4e-67;
Matches 368; Conservative 140; Mismatches 269; Indels 173; Gaps 26;

QY 23 TETVLDSASPDGN-KNGNFSVRESQEDAGTTLFKGNVTLENIPGTGTAITKSCFNNTK 81
Db 26 TNETITSDSYNGNVTSDFEVKET--TSGAIYTCGNVCI-SYAGKDSPLNKSQSFSETT 82
QY 82 GDLTFTGNSLLPOTVDAGTVAGAAVNSVVDKSTTFIGFS--SLSFTASPGSITTTGK 139
Db 83 ENLSFIGNYTLTLCFNTITQSSHPGCAIVSGTNTKLDLSGSLFSCAYCCPG---TTGY 139
QY 140 GAVSCSTGSLTKNVSLFSSKFNSTDNCGAIAKTILSLTGTMTSALFSENTSSKKGAI 199
Db 140 GAIQ-TKGTITLKDNSLSLVFHKNCSTAGGAIQCK----- 173
QY 200 QTSDALTTGNGEVSPSDNTSSDGAALFTEASVTISNNAKVSDFNKVTGASSTTGD 259
Db 174 -----SSSSTAE 180
QY 260 MSGGAICAYKTSTDTKVTLTGNQMLLFSNNTSTTAGGAIYVKLELASGGLTLFNRNSVN 319
Db 181 LK-----LENNKLVFSENSKEKGAIYADKLTIVSGGTLFNSNVS 224
QY 320 GGTAPKGAIAIEDS-GEISLSDSGDGVFLGNTVST---TPGTRSSIDLGTSAKMTA 375
Db 225 HNNSPKGAICIKDSGECSLTANLGDITFDGNKIITNGGSPVTRNSIDLSSGKFTK 284
QY 376 LRSAAGRAIFYDPI-TTGSSTVTDVLKVNETPADSALQYTGNIIFTEGKLSERADSA 434
Db 285 LNAKSGFGIFFDPIANTGGTEI-----ELNKTESDIT--YTKIVFSGEKLSEBKTVP 338
QY 435 KNLTSLKLPVTLSGGTLSLKHGVTLOTAQFTQADSRLEMDVGTTL8-----PADTSTIN 490
Db 339 ANLKSVPKPLKIGAGSLVLDGVTLEAKKITQTKGSTVWMDLGTTLTQTPSSGFTITLT 398
QY 491 NLVINISSI---DGAKKAKIETKATSKNLTLSGTTLLDPTGTFYVENHSLRNQSDVIDEL 547
Db 399 NLDINIASLGGGGTAPAKLATNTASQAISIA-AVNLVNTDSNTYEDPILSASKSFSAI- 456
QY 548 LKASGTVTSTAVTDPPIMGEXF---HYGYQGTWGPVWGTGAST---TATFNWTKTYI 600
Db 457 ---TATTSSTVTPPETNLKNYTPPTPHYQGNW-TVTWKQSSAQEKATLTLWEQTYGYS 512

```

```

QY 601 PNERIGSLVPSLWNAFIDISSLHLYMETANEGLQDRAPFCWAGLSNFFHKKDSTKTRRG 660
Db 513 PNERVGSILVNTLWGAFTSTRAIONLMDISVNGADYSRGFWSSLANFLNKSGDTRKR 572
QY 661 FRHLGGYVIGGNLHTCSDKILSAFQCFGRDRDYFVAKNQGTGYGGLYYQHNETYIS 720
Db 573 FRHSAGYALGVYAQTSPDDVCSAFCQFQKDKDYFVSKNSSTIYAGSIYQH----- 626
QY 721 LPCKLRPCSLSYVPT-----EIPVLPFGNLSYTHDNDLTKY-----TY 761
Db 627 -----ISYMTWNTLQNTLGAEPVLVNAQLTYCHASNMMKMTNTYTPKNVTP 677
QY 762 PTVKSGMGNDSFALEFGGRAPICLDESALPEQYPMFKLOFVYAHQSGFKEQGT-EAREF 820
Db 678 SEIKGDWGNDCFGVEFGAKAPI-ETASLLFDMYSFVKQLQVHAHQDDFKENNSDQGRYF 736
QY 821 GSRRLVNLALPIGIRFDKESDQDATYNLTGLYTVDLVRNPDCTTTLRISGDS--WKTF 878
Db 737 ESNNLTNLSMPIGVKLEKFSHKDTASYNLTLAYAPDIVRNPDPCTASLLVSPTSVAVWTK 796
QY 879 GTNLARQALVLRAGNHCFSNFEAFSOFSELRGSSRNYNVDLGAQYOF 928
Db 797 ANNLAHAFILQAGNYLALTRNTLFSQFELRGSCRTYNIDLGSKIQF 846

RESULT 6
Q8VU49
ID Q8VU49 PRELIMINARY; PRT; 602 AA.
AC Q8VU49
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DE Putative polymorphic membrane protein (Fragment).
OS Chlamydia psittaci (Chlamydia phila psittaci).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phila.
OX NCBI_TaxID=83554;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LLG;
RA Laroucau K., Souriau A., Rodolakis A.;
RT "Isolation of a new pmp sequence and evidence of pmp polymorphism in serotype-1 Chlamydia psittaci strains.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF243418; AAL36962.1; -.
DR InterPro; IPR006315; Autotransport.
DR Pfam; PF02415; DUF145; 1.
DR TIGRFAms; TIGR01414; autotrans_bar1; 1.
FT NON_TER
SQ SEQUENCE 602 AA; 65562 MW; CA486CFACEC131E2 CRC64;

Query Match 25.9%; Score 1237; DB 2; Length 602;
Best Local Similarity 42.4%; Pred. No. 2.5e-50;
Matches 261; Conservative 104; Mismatches 205; Indels 46; Gaps 14;

QY 344 GDIVFLGNTVTSITPGTN---RSSIDLGTSAKMTALRSAGRAIYFYDPI-TTGSSTTVT 399
Db 2 GDITFDGNKIITSRSSSTVKRNSISLGGGKFTKLNAKEGFGIFDYDIANTGDNTEI 61
QY 400 DVLKVNETPADSALQYTNIIFTGKLSLSEADSKNLSKLLQPVTLGGTSLKKGVT 459
Db 62 ELNKAEE---GGSTTYTGKIVFSGEKLSDKEKVADNLKSYFTQPLKIGAGSLVLDGVT 117
QY 460 LQTAQTAQADSRLENDVGTLEPA---DTSTNNLVINISSIDGA---KKAKIETKAT 512
Db 118 LEAKKVSQTDGTVVMDLGTTLQTSSTGETITLNLIDINVASLGGGVGVPADPAKVEAQA 177
QY 513 SKNLTSLGTTLLDPTGTFYENHSLRNPOSYDILELKASGTVTSTAVTPD-----PIMG 566
Db 178 GKTVTIN-AVNLVDTDGNAYEYIPILATSKPFTAIKAGSSGTTTPTDNLKNYTPPT-- 234
QY 567 EKPHYGQGTGPIVNGTGAET---TATFNWTKTGYIPNPERIGSLVPSLWNAFIDISS 623

```

```

Db 235 ---HYGQGNW-TVTKLGTSAQSEETALTWEQTDYSPNERQGPLVNTLMGFSFDIRA 290
QY 624 LHYLMETANEGLQDRAPFCWAGLSNFFHKKDSTKTRRGFRHLGGYVIGGNLHTCSDKILS 683
Db 291 IQNLIDISVNGADYSRGFWVSGGLNHLKSGNSNTRKFRHSAGYALGVYAQTSTEDVFS 350
QY 684 AAFQCFGRDRDYFVAKNQGTGYGGLYYQHNETYISLPCLRPCSLSYVPTIPLVLSFG 743
Db 351 AAFQCFGRDRDYFVAKNQGTGYGGLYYQHNETYISLPCLRPCSLSYVPTIPLVLSFG 743
QY 744 NLSYTHDNDLTKY-----TYPTVKGSGMGNDSFALEFGGRAPICLDESALPEQYPM 795
Db 407 QLYYCHASNMMKMTNTYTPKNVTLSEIKGDWGNDCFGVEFGAMAPIETPSSFLFDYS 466
QY 796 PFMKLOFVYAHQSGFKEQGT-EAREFGSSRLVNLALPIGIRFDKESDQDATYNLTGLYT 854
Db 467 PFLQQLVHAHQDDFKENNSDQGRYFESSNLTNLSMPIGIVKPERFAVNDVASYHLTAAYA 526
QY 855 VDLVRNPDCTTTLRISGDS--WKTFGTNLARQALVLRAGNHCFSNFEAFSOFSELR 912
Db 527 PDIVRNPDPCTASLLVSPTSVAVWTKANNLARSFAFMQLQAGNYLALSHNMELFSQGF 586
QY 913 GSSRNYNVDLGAQYOF 928
Db 587 GSSRNYNVDLGSKIQF 602

RESULT 7
Q8VL57
ID Q8VL57 PRELIMINARY; PRT; 601 AA.
AC Q8VL57
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DE Putative polymorphic membrane protein (Fragment).
OS Chlamydia psittaci (Chlamydia phila psittaci).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phila.
OX NCBI_TaxID=83554;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POS, and LLG;
RA Laroucau K., Souriau A., Rodolakis A.;
RT "Isolation of a new pmp sequence and evidence of pmp polymorphism in serotype-1 Chlamydia psittaci strains.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF243417; AAL36961.1; -.
DR InterPro; IPR006315; Autotransport.
DR Pfam; PF02415; DUF145; 1.
DR TIGRFAms; TIGR01414; autotrans_bar1; 1.
FT NON_TER
SQ SEQUENCE 601 AA; 65476 MW; D6AA97EC9072C757 CRC64;

Query Match 25.5%; Score 1219.5; DB 2; Length 601;
Best Local Similarity 42.0%; Pred. No. 1.6e-49;
Matches 259; Conservative 106; Mismatches 204; Indels 47; Gaps 15;

QY 344 GDIVFLGNTVTSITPGTN---RSSIDLGTSAKMTALRSAGRAIYFYDPI-TTGSSTTVT 399
Db 2 GDITFDGNKIITSRSSSTVKRNSISLGGGKFTKLNAKEGFGIFDYDIANTGDNTEI 61
QY 400 DVLKVNETPADSALQYTNIIFTGKLSLSEADSKNLSKLLQPVTLGGTSLKKGVT 459
Db 62 ELNKAEE---GGSTTYTGKIVFSGEKLSDKEKVADNLKSYFTQPLKIGAGSLVLDGVT 117
QY 460 LQTAQTAQADSRLENDVGTLEPA---DTSTNNLVINISSIDGA---KKAKIETKAT 512
Db 118 LEAKKVSQTDGTVVMDLGTTLQTSSTGETITLNLIDINVASLGGGVGVPADPAKVEAQA 177
QY 513 SKNLTSLGTTLLDPTGTFYENHSLRNPOSYDILELKASGTVTSTAVTPD-----PIMG 566
Db 178 GKTVTIN-AVNLVDTDGNAYEYIPILATSKPFTAIKAGSSGTTTPTDNLKNYTPPT-- 234

```

```

QY 567 EKHYGQGTWGPVWGTGAST--TATFNWTKTYIIPNPRIGSLVPSNLWNAFIDISS 623
D 235 ---HYGQGNW-TVTWKLGTSAQBETATLTWEQTDYSPNPERQGPLVPTKLWGSFSDIRA 290
QY 624 LHYLMETANEGLQGDRAFCWAGLNFPHKDSKTKTTRGPHLSGGVVGNNLHSCDKILS 683
D 291 IQNLIDISVNGADYRGRFWWSGLNLFHKSQNTKRKRPHHSAGYALGVYAQTSTEDVFS 350
QY 684 AAFQCLQGRDRDYVAKNQGTGYGTTLYYQHNETYISLPCKLRPCSLSYVPTPIVLFSG 743
D 351 AAFQCLQFKDKDYFVKNSNNIYAGSIYYQHSYWNWQNLQ---STICAEAPLVNA 406
QY 744 NLSVTHDNDLTKYIT-----TYPVKGSGWNSDPALEFGGRAPICLDESALFQYM 795
D 407 QLTCHASNNMKNMTNTYVKNVTLSEIKGWDGNCDFGVFGAMAPI-ENASFLFDYIS 465
QY 796 PFMKLQFYVAHQEGFKQGT-EAREFGSSRLVNLALPIGIRFDKESQCDQDATYNTLGYT 854
D 466 PFLQQLVHAHQDDPKENNSQGRYFESSNLTNLSMPIGIKFERPAYNDVASYHLTRAAYA 525
QY 855 VDLVRSPDCPTTLRLISGDS--WKTFTGNLQALVLRAGNHFCFNSNFEAFSOFSELR 912
D 526 PDIVRSNPDCTASLLVSPSAVWTKANLARSAPMLQAGNVYALSHNMELFSQFGPEIR 585
QY 913 GSSRNYNVDLQAKYQF 928
D 586 GSSRTYNVDLGSKIQF 601

RESULT 8
Q8VU50 PRELIMINARY; PRT; 700 AA.
AC Q8VU50;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative polymorphic membrane protein (Fragment).
OS Chlamydia psittaci (Chlamydia phila psittaci).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phila.
OX NCBI_TaxID=83554;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POS;
RA Laroucau K., Souriau A., Rodolakis A.;
RT "Isolation of a new pmp sequence and evidence of pmp polymorphism in
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF243415; AAL36959.1; -.
DR InterPro; IPR003368; Chlamydia_PMP.
DR TrEMBL; AF243415; AAL36959.1; -.
DR TrEMBL; AF243415; AAL36959.1; -.
FT NON TER 700 700
SQ SEQUENCE 700 AA; 74200 MW; 2C38E504AF48D1D1 CRC64;

Query Match 24.6%; Score 1175; DB 2; Length 700;
Best Local Similarity 35.5%; Pred. No. 2.4e-47;
Matches 284; Conservative 120; Mismatches 227; Indels 170; Gaps 25;

QY 23 TETVLDSASPOGN-KNGNFSVRESQEDAGTTLFKGNVLTENIPGTGTAITKSCFNNTK 81
D 26 TNETLSSSYNGVNTSDDFEVKET--TSGLIYTCGNVCI-SYAGKSPINKSCFSETT 82
QY 82 GDLTFTGNGNSLLFQVWDAGTVAGAAVNSVVDKSTFTFGFSLSFIAFSGSIITGKA 141
D 83 ENLSFIGNGYTLCFDNIITTAGNPGAINVSGQKTLNVSGLFSFCAHCPGP--TTGYGA 140
QY 142 VSCSTGSLTLKNVSLLSKFNSTDNGAITAKTILSTGTTMSALFSENSTSSKKGAIQT 201
D 141 IQ-----TKGV-----
QY 202 SDALATITGNOGEVFSFSDNTSDSGAAIFTEASVTISNNAKVFIQDKVTKGSSSTTGDM 261

```

```

D 148 -----TPSGNNKLIIDNNCSTG-----E 165
QY 262 GGAI-CAYKSTDTKVTLTQNMQLFSNNTSTTAGGAIYVKLELAGGGTLFSPNSVNG 320
D 166 GGAIKCA--TGSNAELKLEGNYSVVFSGNSQKGGAIYTKLITADGPTLFSNNSVS- 222
QY 321 GTAPKGAIAIED-SGELSLSADSGDIVLGNVTSTTPGNT--RSSIDLGTSAKMTAL 376
D 223 ASSPKGGAICUDDTSSECSLUNALGDITFGNKVITNGSSSTVKRAINDLGGGKTKL 282
QY 377 RSAAGRAIYFDPYI-TTSSSTTVDVLKVNNETPADSALOYTGNIIIFTGKLSSETAADSK 435
D 283 NAKGPGCIPFHDPIANTGGSTEI-----ELNKTESDTT--YTGKIVFSGEKLSDBEKTVDP 336
QY 436 NLTSKLLQPTVTLGGTSLKHGVTLTQOAFQOADSRLMDVGGTTLB-----PADTSTINN 491
D 337 NLKSYFQPLKIGAGSLVLDKGVTLKAKITQTKGSTVVMVMDLGTTLQTPSSSGETITLTN 396
QY 492 LVINISSI--DGAKKAKIETKATSKNLITLGGITLADPTGFVENHSLRNPOSYDILEL 548
D 397 LDINIASLGGGGGTAPAKLATNTASQAISIA-AVNLVNTSDSYVEDPILSASKSFSAI-- 453
QY 549 KASGTVTSTAVTDPPIMGKEF---HYGYQGTWGPVWGTGAS---TATFNWTKTYIIP 601
D 454 --TATTSSSTVTPPETNLKNVTPPTHYGYQGNM-TLAWPPGETMQLKTATLNWEQTCYSP 510
QY 602 NPERIGSLVPSNLWNAFIDISSLHYLMETANEGLQGDRAFCWAGLNFPHKDSKTKTRGF 661
D 511 NPERVGSVPLNTLWGSFSDIRAIQNLMDVSNGADYSRGFWVSGLANFLNKGSDTKRKF 570
QY 662 RHLSGGVVGNNLHSCDKILSAAFCOLFGRRDRDYFVAKNQGTGYGTTLYYQHNETYISL 721
D 571 RHNSAGVALGVYAQTSPEDIFSAACOLFGKDKDYFLSKNSSTIYAGSIYYQH----- 623
QY 722 PKLRLPCLSY-----VPTPIVLFSGNSLSTHTDNDLTKYT-----TYP 762
D 624 -----ISYWNWQNLQNTICAEAPLVNAQLTYCHASNNMKNMTNTYTPKNVTPS 675
QY 763 TVKSGWNSDPALEFGGRAPY 783
D 676 EIKGWSGDCFCGVFGAKAPI 696

RESULT 9
Q8VU48 PRELIMINARY; PRT; 581 AA.
AC Q8VU48;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Putative polymorphic membrane protein (Fragment).
OS Chlamydia psittaci (Chlamydia phila psittaci).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phila.
OX NCBI_TaxID=83554;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POS;
RA Laroucau K., Souriau A., Rodolakis A.;
RT "Isolation of a new pmp sequence and evidence of pmp polymorphism in
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF243415; AAL36963.1; -.
DR InterPro; IPR003368; Chlamydia_PMP.
DR TrEMBL; AF243415; AAL36963.1; -.
FT NON TER 581
SQ SEQUENCE 581 AA; 62860 MW; CDDF3C98522E112F CRC64;

Query Match 24.6%; Score 1172.5; DB 2; Length 581;
Best Local Similarity 40.4%; Pred. No. 2.5e-47;
Matches 253; Conservative 100; Mismatches 185; Indels 89; Gaps 17;

QY 344 GDIVFLGNVTSTTPGNT--RSSIDLGTSAKMTALRSAGRAIYFYDPI-TTGSSTTVT 399

```

```

Db      2 GDITFDGNKILITTSRSSSTVKRNSISLGGGKFTKLNAKEGFGIFFYDPIANTGDTNTEI 61
Qy      400 DVLUKNETPADSALOYTNIIFTGEKLSSETAADSKNLTKLQPVLTSGGTSLSKHGVT 459
Db      62 ELNKAEE---GGSTYTGKIVFSEKLSDBEKKVADNLKSVFTQPLKIGAGSLVLDGVT 117
Qy      460 LQOQAFQQAADSRLEMDVGTLEPA---DSTINNLVINISSIDGA---KKAKIETKAT 512
Db      118 LEAKKVQQTGSTAVMDLGTLOTSSSGEITLTNLNDINVASLGGGGAADPAKVEAQAAS 177
Qy      513 SKNLTLSGTITLDPDTGTFYENHSLRNPQSDIILELKASGTVTSTAVTPD-----PIMG 566
Db      178 KKTVIN-ANVLVDTDGNAEYPIATSQPTAIIAKAGSGTITTTFDLKNVTPPT-- 234
Qy      567 KEHYGQGTGWPVWGTGAST---TATFNWTKGYIPNPERIGSLVPSNLWNAFIDISS 623
Db      235 ---HYGQGNW-TVTWKLGTSAQEBTATLNWEQGTGYSFNPBRVGSLSVPLTNGSFSDIRA 290
Qy      624 LHVLMETANEGLODRAFWCAGLSNPFHKDSTKTRRGFRHLHSGGVYVGGNHLTCSDKILS 683
Db      291 IQNLMDVSVNGADYSRGFWGSLANFLNKGSGDTKRRFRHNSAGYALGVYAQTPESDIFS 350
Qy      684 AAFQQLFGRDRDYFVAKNOGTGVYGGTLYYQHNETYIISLPCKLRPCSLY----- 732
Db      351 AAFQQLFGKDKYFLSKNSSTIYAGSIYQH-----ISYNNWQNLLQNT 395
Qy      733 VPTEIPVLFSGNLSYTHTDNDLTKYTT-----TYPTVKGSGWNSDFALEFGGRAPIC 784
Db      396 IGAEAPVLVLAQTYCHASNMTNTMTVTPKNVTPSEIKGWDGSCFCGVEFGAKAPI- 454
Qy      785 LDESALFEQWPMKLFQVVAHQEGEKEQGT-EAREFGSRLSNLALPIGIRDKESDCQ 843
Db      455 ETASLLPDMTSPFVKVLQVLAHQDQFKENNSDQGYFESNNLNLWSMPIGVKLEKFSHED 514
Qy      844 DATYNTLTGYTVDLVRNPDCTTTLRISGDS--WKTFGTNLARQALVLRAGNHFCFNSNF 901
Db      515 TASVNLTLAVAPDIVRNPDPCTASLLVSPTSAVWVTKANLARHAFILOAGNYLA----- 569
Qy      902 EAFSQFSELRGSSRNYNVDLGAKYQF 928
Db      570 -----TYNIDLGSKIQF 581

RESULT 10
Q9RB67 ID Q9RB67 PRELIMINARY; PRT; 445 AA.
AC Q9RB67;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Pmp_5.
DE Pmp_5.2.
GN Pmp_5.2.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiales; Chlamydiales; Chlamydiales.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Iehli K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RL Nucleic Acids Res. 28:2311-2314(2000).
DR EMBL; AP002545; BAA98231.1;
DR InterPro; IPR006315; Autotransport.
DR InterPro; IPR005546; Autotransporter.
DR Pfam; PF03797; Autotransporter; 1.
DR TIGRFAMs; TIGR01414; autotrans_bar1; 1.
SQ SEQUENCE 445 AA; 49353 MW; EFA24AFC9C5097A6 CRC64;

Query Match 21.3%; Score 1017.5; DB 16; Length 445;
Best Local Similarity 45.3%; Pred. No. 3.2e-40;

Matches 202; Conservative 70; Mismatches 165; Indels 9; Gaps 7;

Qy 489 INNLVINISSIDGAKAKIETKATSKNLTLGSGTITLDPDTGTFYENHSLRNPQSDIILEL 548
Db 3 ITNLSINADITIKNPNIVASAAKNITLTGTLALVNDAGAYENHTLQDSQDYSFVKL 62
Qy 549 K--ASGTVTSTAVTPD--IMGEKPHYGYQGTWG-PIVWGTGAS-TTATFNWTKGYIPN 602
Db 63 SPGAGGTITITQDASQKPLEVAPSRPHYGYQHNWVQVPGTGTQPSQANLEWRTGYLPN 122
Qy 603 PERIGSLVPSNLWNAFIDISSLHLMETANEGLODRAFWCAGLSNPFHKDSTKTRRGFR 662
Db 123 PERQGSGLVPSNLWNGSFVDQRAIOEIWNSSQILCQERGWWAGIANFLHRDKI-NBHGYS 181
Qy 663 HLSGGYVIGGNLTCSDKILSAFAFCQLFGRDRDYFVAKNOGTGVYGGTLYYQHNETYISLP 722
Db 182 HSGVGLVGVGTHTAFSDATINAAFCQLFGRDKDYVVKNGHTSYSGVVFLEDTLEFRS-P 240
Qy 723 CKLRPCSLSVPTPEIPVLFSGNLSYTHTDNDLTKYTTPTVKGSGWNSDFALEFGGRAP 782
Db 241 QGFYTDSSSEACCNQVVTIDMQLSYSHRNDMKTITTYPEAQGSWANDVFGLEFGATTY 300
Qy 783 ICLEDSALPEQWPMKLFQVVAHQEGEKEQGT-EAREFGSRLSNLALPIGIRDKESDC 842
Db 301 YYPNSTFLFDYSPFRLQCTYAHQEDFKETGGEVRHFTSGDLFLNLAVPVIGVKFERFSDC 360
Qy 843 QDATYNTLTGYTVDLVRNPDCTTTLRISGDSWKTFGTNLARQALVLRAGNHFCFNSNF 902
Db 361 KRGSYELTPAYVPDIVRKDPKSTATL-ASGATWSTHGNLSRQGLQLRGNHCLINPGIE 419
Qy 903 AFSQFSELRGSSRNYNVDLGAKYQF 928
Db 420 VFSHGATELRGSSRNYNINLGKRYF 445

RESULT 11
P71134 ID P71134 PRELIMINARY; PRT; 649 AA.
AC P71134;
DT 01-FEB-1997 (T-EMBLrel. 02, Created)
DT 01-FEB-1997 (T-EMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Putative outer membrane protein (Fragment).
OS Chlamydia abortus.
OC Bacteria; Chlamydiales; Chlamydiales; Chlamydiales; Chlamydiales.
OX NCBI_TaxID=83555;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ovine abortion S26/3;
RA Longbottom D., Russell M., Dunbar S.M., Jones G.E., Herring A.J.;
RT "98kDa protein genes from ovine abortion strain S26/3 Chlamydia
RT psittaci.";
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U72499; AAB18187.1;
DR InterPro; IPR006315; Autotransport.
DR InterPro; IPR003388; Chlamydia_PMP.
DR Pfam; PF02415; DUF145; 1.
DR TIGRFAMs; TIGR01414; autotrans_bar1; 1.
FT NON TER 1
SQ SEQUENCE 649 AA; 70091 MW; 13747C68066A7F50 CRC64;

Query Match 19.6%; Score 935.5; DB 2; Length 649;
Best Local Similarity 31.5%; Pred. No. 3.5e-36;
Matches 217; Conservative 103; Mismatches 243; Indels 125; Gaps 12;

Qy 322 TAPKGAIAIEDSGELSLGADSDIIVFLNGTITVSTTPGNNRISIDLGTSKMTALSAAG 381
Db 6 TAPAAAAVVTAKOPEASLAKATSGIPASGASVSPAPAPTPTKPLK----- 51
Qy 382 RAIYFYDDPITGSGTSTVTDVLLKNVETPADSALQYTCNIIFTGKLTSETAADSKNLTSKL 441
Db 52 -----NAPDTPDPIQKVAEAAQGSVYNGKIVFSGEKLSEDAKNPLNATSVI 101

```



```
Qy 255 STTGMS-----GGAICAYKTTDTKVTILTCNQMLLFNSNTSTTAGGA 297
Db 258 NCSGDLTDTNTSLLLOENSTMQGGALC-----STGT-ISTGSDSINVIGNTSGQKGA 312
Qy 298 IYVKKLELASG-GLTLFSRNSVNGGTAPKGAIAIEDSGELSLSDSGDIVFLGNTVTST 356
Db 313 ISASLXILGGQGGALFSNNVVTHAT-PLGGAIFINTGGSILQFTQGGDIVFEGNQVTTT 371
Qy 357 TPG--TNRSSIDLGTSAKMTALRSAAGRAIYFYDPITTTGSSSTTVDVLKNVETPADSALQ 414
Db 372 APNATTXRNVIHLSSTAKWTGLAASQGNAIYFYDPITT-NDTGASDNLRLINEVSANOKL- 429
Qy 415 YTGNIITGKLSSTAADSKNLTSLKLOPVTLTSGGTLISKHGYTLQTOFTQQAADSRLE 474
Db 430 -SGSIVFSGERLSTAEAT-AENLTSRINQPVTLVEGSLVLKQGVTLITQGSQBFESTLL 487
Qy 475 MDVGTTL 481
Db 488 LDLGTSL 494

RESULT 14
Q9RB70 PRELIMINARY; PRT; 427 AA.
AC Q9RB70;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Polymorphic outer membrane protein G family.
GN PMP_4.1.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
DR EMBL; AP002545; BAA98227.1; -
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF02415; DUF145; 1.
DR TIGRFAMs; TIGR01376; POMP_repeat; 6.
SQ SEQUENCE 427 AA; 43419 MW; AB4BBEC1594DD2B1 CRC64;

Query Match 16.6%; Score 792; DB 16; Length 427;
Best Local Similarity 44.0%; Pred. No. 1.1e-29;
Matches 198; Conservative 70; Mismatches 146; Indels 36; Gaps 14;

Qy 1 MKSFPKPFVFTFAIFPLSMI-----ATETVLDSSASFDGN-KNGNFVSRESQEDA-GTTY 54
Db 1 MRSFSLLLIISSLAFFPLMSVSDAADLTLSRDSYNGDTSTTEFTPKAATSDASGTTY 60
Qy 55 LFGKNTLENIPTGCTAITKSCFNNTKGDLTFTGNGSNLLFQTVDAAGTVAAGAAVNSVVD 114
Db 61 ILDGDVSIQ-AGKQTSLTSCFSNTAGNLATFLNGFSLHFDNIISSTVAGVWVNTAAS 119
Qy 115 KSTTFIGPSSLSFTASPGSSITTKGAVSCSTGSLTKVNSLLFKNPFSTDNGCAITAK 174
Db 120 GITKFGSFTLRMLAAPR---TTGKAIKITDG-LVFESIGNLDLNENASSENGCAINTK 175
Qy 175 TLSLTGTMSALFSENSTSKKGGAIQTSDALITIGNQGEVSFSDNTSSDGAALFTBASV 234
Db 176 TLSLTGSTRFVAFNLGSSSQGGALYASGDSVISENAGILSFGNNSATTSAGIAEAGNL 235
Qy 235 TISNNKVSFTDKNVKTGASSSTTGMSCGACAYK--TSTDTKVTLTNQMLLFNSNTST 292
Db 236 VISNNQNIFFDGCRAI-----TNGGAIDCNKAGANPDPIILTLSGNESLHFLNNTAG 286
```

```
Qy 293 TAGAIYVKKLELASG-GLTLFSRNSVNGGTAPKGAIAIEDSGELSLSDSGDIVFLGN 351
Db 287 NSGAIYTKLVLSRGGVLFNNKAAT-PAKGAIALDSGEISISADLGNLIFEGN 345
Qy 352 TVTSTT---PGTNRSSIDLGTSAKMTALRSAAGRAIYFYDPITTTGSSSTTVDVLKNVETP 408
Db 346 T-TSTTGSPASVTRNAIDLASNAKFLNLRATRGKVIYFDPIT---SSGATDKLSLNKAD 401
Qy 409 ADSALQYTGNIITGKLSSTAADSKNLT 438
Db 402 AGSGNTYEGYIVFSGEKLSEV-----RNLT 426

RESULT 15
Q9RB72 PRELIMINARY; PRT; 186 AA.
AC Q9RB72;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Polymorphic outer membrane protein G family.
GN PMP_3.1.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
DR EMBL; AP002545; BAA98225.1; -
DR InterPro; IPR003368; Chlamydia_PMP.
DR Pfam; PF02415; DUF145; 1.
DR TIGRFAMs; TIGR01376; POMP_repeat; 6.
SQ SEQUENCE 186 AA; 19540 MW; 64557A5346533FA3 CRC64;

Query Match 16.1%; Score 768; DB 16; Length 186;
Best Local Similarity 95.1%; Pred. No. 5.1e-29;
Matches 155; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MKSFPKPFVFTFAIFPLSMIATETVLDSSASFDGNKNGNFVSRESQEDAGTTVLFKGNV 60
Db 10 MKSFPKPFVFTFAIFPLSMIATETVLDSSASFDGNKNGNFVSRESQEDAGTTVLFKGNV 69
Qy 61 TLENIPTGCTAITKSCFNNTKGDLTFTGNGSNLLFQTVDAAGTVAAGAAVNSVVDKSTTFI 120
Db 70 TLENIPTGCTAITKSCFNNTKGDLTFTGNGSNLLFQTVDAAGTVAAGAAVNSVVDKSTTFI 129
Qy 121 GFSSLSFIASPGSSITTKGAVSCSTGSLTKVNSLLFKNPF 163
Db 130 GFSSLSFIASPGSSITTKGAVSCSTGSLTKVNSVCSAKTF 172

Search completed: December 16, 2003, 10:54:02
Job time : 47 secs
```

